CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/71, 14/495, 16/28, 16/42, G01N 33/68, A61K 38/17, 39/395, C12N 5/10

A1

(11) International Publication Number: WO 95/10610

(43) International Publication Date: 20 April 1995 (20.04.95)

(21) International Application Number: PCT/US94/11648

(22) International Filing Date: 14 October 1994 (14.10.94)

(30) Priority Data:
08/138,368 15 October 1993 (15.10.93) US

(71) Applicant: LA JOLLA CANCER RESEARCH FOUNDATION [US/US]; 10901 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventors: RUOSLAHTI, Erkki; P.O. Box 1054, Rancho Santa Fe, CA 92067 (US). FUKUSHIMA, Daikichi; 2-99-21-403, Oharano, Nishi Kyoku, Kyoto 610-11 (JP). BUTZOW, Ralf; Toolonkatu 7A8, FIN-00100 Helsinki (FI).

(74) Agents: CAMPBELL, Cathryn et al.; Campbell and Flores, Suite 700, 4370 La Jolla Village Drive, San Diego, CA 92122 (US).

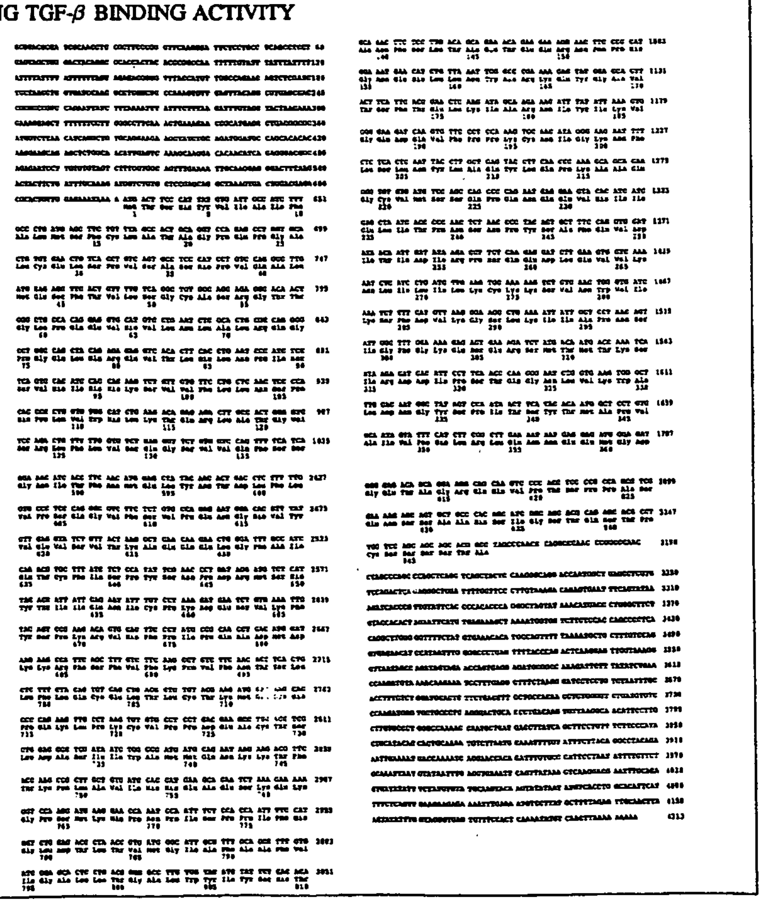
(81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, FI, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LV, MD, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

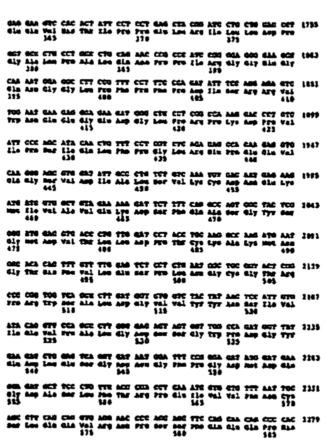
Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: BETAGLYCAN POLYPEPTIDES HAVING TGF-β BINDING ACTIVITY

(57) Abstract

This invention provides polypeptides that bind to TGF-β and that have an amino acid sequence of a portion of a mammalian betaglycan. These polypeptides enhance TGF-β binding to type II receptor and enchance the suppression of cell growth by TGF-β. This invention also provides isolated nucleic acid molecules encoding these polypeptides. This invention also provides methods of detecting and isolating TGF-β. This invention also provides pharmaceutical compositions having these peptides and therapeutic uses of them. This invention also provides decoy betaglycan polypeptides, anti-betaglycan-binding-site antibodies and anti-idiotypic antibodies against anti-betaglycan. It also provides uses of these compositions.





* (Referred to in PCT Gazette No. 26/1995, Section II)

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

BETAGLYCAN POLYPEPTIDES HAVING TGF-β BINDING ACTIVITY

BACKGROUND OF THE INVENTION

This invention relates to the field of polypeptides and, in particular, to betaglycan polypeptides that bind to TGF-β.

This invention was made partially with government support under CA 42507 and CA 30119 awarded by the National Cancer Institute. The government has certain rights in this invention.

Transforming growth factor-beta ("TGF-ß") is a multifunctional cytokine that plays an important role in regulating repair and regeneration following tissue injury. Three isoforms of TGF-ß (TGF-ß1, 2 and 3) are expressed in mammals and to date show similar properties *in vitro*. Platelets contain high concentrations of TGF-ß, and upon degranulation at a site of injury, release TGF-ß into the surrounding tissue. TGF-ß then initiates a sequence of events that promotes healing including (1) chemo-attraction of monocytes, neutrophils, and fibroblasts, (2) auto-induction of TGF-ß production and stimulation of monocytes to secrete interleukin-1 (IL-1), tumor necrosis factor and other cytokines, (3) induction of angiogenesis and cell proliferation, (4) control of inflammation and cell toxicity by acting as a potent immunosuppressant and inhibitor of peroxide release, and (5) increased deposition of extracellular matrix. TGF-ß also induces proliferation of macrophages exposed in combination with macrophage colony stimulating factor ("M-CSF") or granulocyte macrophage colony stimulating factor ("GM-CSF").

However, the excessive action of TGF-β is associated with pathological scarring and has been implicated in glomerulonephritis, diabetic nephropathy, lung fibrosis, liver cirrhosis, intimal hyperplasia,

cardiac cirrhosis after infarct, adult respiratory distress syndrome and other fibrotic pathologies.

Thus, the regulation of TGF-ß activity, both up and down, has important therapeutic significance.

At least three proteins have been identified as part of the cell surface receptor system that transduces the signal from TGF-β to the cell: type I receptor (RI), type II receptor (RII), and betaglycan (type III receptor). Betaglycan is a transmembrane proteoglycan whose core protein has molecular weight of about 100 kD. The betaglycan core protein binds TGF-ß with high affinity.

The betaglycan core protein has 853 amino acids and consists of four domains. The N-terminal one-third of the extracellular domain and the C-terminal cytoplasmic domain have similarities with endoglin. Endoglin is a TGF-β binding protein expressed on the surface of endothelial cells. The middle part of the betaglycan extracellular domain has a short stretch similar to a portion of endoglin but otherwise bears no homology to any known protein. The 260 amino acids in the ectodomain closest to the membrane are related to a domain in a group of transmembrane proteins which include sperm receptors Zp2 and Zp3, the zymogen granule membrane protein, GP2, and uromodulin. This common domain occurs at a similar location relative to the transmembrane domain in these proteins.

Betaglycan is the major TGF-ß-binding molecule on most cell types. However, many cell types which respond to TGF-ß, e.g., hematopoietic cells, do not appear to have detectable amounts of betaglycan. This, together with evidence from TGF-ß-resistant mutant cell lines, suggests that betaglycan is not directly involved

in the TGF-ß signal transduction pathway. Over-expression of betaglycan in Chinese hamster ovary cells enhances binding of TGF-ß to type II receptor, suggesting that betaglycan acts by stockpiling TGF-ß and presenting it to the signal transducing proteins in the receptor.

Thus, betaglycan has potential as a modulator of TGF-ß bioactivity. However, the art does not disclose the portion of betaglycan that has TGF-β binding activity and that enhances TGF-β binding to the type II receptor. This invention satisfies this need and provides additional advantages by identifying a portion of betaglycan that binds TGF-β, enhances TGF-β binding to the type II receptor and enhances suppression of cell growth by TGF-β.

SUMMARY OF THE INVENTION

This invention provides polypeptides of at least 155 amino acids that bind to TGF-β and that have a sequence consisting essentially of a sequence of a portion of a mammalian betaglycan within about one-third of the extracellular domain closest to the cell membrane. It also provides polypeptides having a sequence consisting essentially of a portion of a mammalian betaglycan wherein the portion is about one-fourth or about one-fifth of the extracellular domain of a mammalian betaglycan closest to the cell membrane. More particularly, it provides polypeptides wherein the sequence consists essentially of at least amino acids 543 to 769 of SEQ ID NO:2 to at most amino acids 501 to 853 of SEQ ID NO:2.

This invention also provides a soluble polypeptide having the formula A-B-C, wherein A is a sequence excluding amino acid sequences of more than 4 amino acids from amino acids 1 to 543 of SEQ ID NO:2; B

is an amino acid sequence consisting essentially of at least amino acids 543 to 769 of SEQ ID NO:2 to at most amino acids 501 to 769 of SEQ ID NO:2; and C is an amino acid sequence.

This invention also provides a soluble polypeptide having the formula A-B-C, wherein A is an amino acid sequence excluding sequences of more than 4 amino acids from amino acids 1 to 543 of SEQ ID NO 2; B is at least 155 amino acids in a sequence consisting essentially of an amino acid sequence within amino acids 543 to 769 of SEQ ID NO:2; and C is an amino acid sequence.

This invention also provides isolated nucleic acid molecules encoding any of the polypeptides of this invention, expression vectors having an expression control sequence operatively linked to a nucleic acid molecule of this invention, and prokaryotic or eukaryotic cells transfected with an expression vector of this invention and capable of expressing a nucleic acid of this invention.

This invention also provides methods of detecting TGF-β in a sample by contacting the sample with a polypeptide of this invention and determining the amount of TGF-β bound to the polypeptide.

This invention also provides methods of isolating TGF-β from a sample by contacting the sample with a polypeptide of this invention bound to a solid support to allow binding of TGF-β to the polypeptide and isolating the TGF-β from the polypeptide.

This invention also provides methods of enhancing the binding of TGF-ß to a TGF-β receptor by

contacting a cell bearing a TGF-β receptor with TGF-β and a polypeptide of this invention.

This invention also provides methods of enhancing suppression of cell growth by TGF-β comprising contacting a cell with TGF-β and a polypeptide of this invention.

This invention also provides pharmaceutical compositions comprising a polypeptide of this invention in a pharmaceutically acceptable carrier.

This invention also provides methods of treating a subject with a condition ameliorated by the enhanced binding of TGF-β to a TGF-β receptor or by suppression of cell growth by TGF-β by administering a therapeutically effective amount of a pharmaceutical composition having a polypeptide of this invention.

This invention also provides decoy betaglycan polypeptides. It also provides methods of treating a subject with a condition ameliorated by the diminished binding of TGF-β to a TGF-β receptor or by the inhibition of the suppression of cell growth by TGF-β by administering to the subject a therapeutically effective amount of a pharmaceutical composition having a decoy betaglycan polypeptide. It also provides methods of suppressing TGF-β-induced deposition of extracellular matrix in a subject by administering to the subject a therapeutically effective amount of a pharmaceutical composition having a decoy betaglycan polypeptide.

This invention also provides anti-betaglycan-binding-site antibodies. It also provides methods of treating a subject with a condition ameliorated by the diminished binding of TGF-β to a TGF-β receptor or by the inhibition of the suppression of cell growth by TGF-β by

administering to the subject a therapeutically effective amount of a pharmaceutical composition having an anti-betaglycan-binding-site antibody of this invention. It further provides methods of suppressing TGF-β-induced deposition of extracellular matrix in a subject by administering to the subject a therapeutically effective amount of a pharmaceutical composition having an anti-betaglycan-binding-site antibody of this invention.

This invention also provides anti-idiotypic antibodies that mimic the ability of a polypeptide of this invention to bind TGF-β but that lack the biological function of enhancing binding of TGF-β to the type II receptor or enhancing suppression of cell growth by TGF-β. It also provides methods of treating a subject with a condition ameliorated by the diminished binding of TGF-β to a TGF-β receptor or by inhibition of the suppression of cell growth by TGF-β by administering to the subject a therapeutically effective amount of a pharmaceutical composition having an anti-idiotypic antibody of this invention. It also provides methods of suppressing TGF-β-induced deposition of extracellular matrix in a subject by administering to the subject a therapeutically effective amount of a pharmaceutical composition having an anti-idiotypic antibody of this invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B depict the coding region of rat betaglycan cDNA. The filled part is the transmembrane region. Restriction sites used to generate the cDNA fragments are shown. The brackets show the cDNA fragments that were expressed as proteins (bg1 through bg4). Figure 1B depicts recombinant fragments of betaglycan analyzed by SDS-PAGE (4 % to 20 % gels) and Coomassie Blue staining.

Figures 2A and 2B present results of a competition assay for the binding of ^{125}I-TGF-ß1 to Hep G2 cells by betaglycan fragments. Hep G2 cells on 24-well dishes were incubated with ^{125}I-TGF-ß1 (100 pM) and with indicated concentrations of betaglycan fragments that had been expressed as fusion protein in bacteria. The bound TGF-ß was measured, and the results are expressed as a percent of total binding. Vertical bars show the S.D. of triplicate samples. T=total binding of ^{125}I-TGF-ß1, NS = nonspecific binding measured as residual binding in the presence of 20 nM of unlabeled TGF-ß1. Figure 2A presents the effect of fragments bg1, bg2, bg3, and bg4 and sub-fragments of bg3. Figure 2B presents the effect of bg3 and combinations of bg3 sub-fragments.

Figure 3 shows gel electrophoresis of bg3 betaglycan fragments on the binding of ^{125}I-TGF-ß to MV1 Lu cells in receptor affinity labeling. MV1 Lu cells on 6-well dishes were affinity labeled by using 100 pM of ^{125}I-TGF-ß1 at 37°C without or with various concentrations of the bg3 fragment. After cross-linking the cells were solubilized and analyzed in SDS-PAGE followed by autoradiography. The band at 71 kDa is RI, the bands at and below the 101 kDa marker represent RII and the wide band above the 208 kDa is betaglycan.

Figures 4A and 4B present results of binding studies of the bg3 betaglycan fragment to TGF-ß1 in a solid-phase binding assay. In Figure 4A, ^{125}I-labeled bg3 fragment was incubated in microtiter wells coated with increasing concentrations of TGF-ß1 in the buffer (●) and in the presence of unlabeled bg3 fragment (○) (1μM). In Figure 4B, ^{125}I-bg3 was incubated for the times indicated in microtiter wells coated with 1 μg/ml of TGF-ß1. Binding is expressed as percent of the radioactivity added to the wells. Error bars show standard deviation of triplicate samples.

Figure 5A and 5B show competition and Scatchard plots for the binding of ^{125}I-bg3 betaglycan fragment to immobilized TGF-ß1. Figure 5A shows competition of ^{125}I-bg3 binding to TGF-ß1 by increasing concentrations of bg3 fragment. Figure 5B shows Scatchard plot of the binding data in Figure 5A.

Figure 6 shows results of a competition assay of the binding of ^{125}I-bg3 betaglycan fragment to immobilized TGF-ß1 by core proteins of decorin-type proteoglycans. ^{125}I-bg3 was incubated in microtiter wells coated with TGF-ß1 (1μg/ml) in the presence of various concentrations of unlabeled bg3 fragment (○) or fusion proteins of human decorin (■), human biglycan (▲), and human fibromodulin (●) core proteins, or maltose-binding protein (♦), the protein the proteoglycan core proteins were fused to. Binding is expressed as a percent of binding of the ^{125}I-bg3 in the absence of competitors.

Figure 7 shows the effect of bg3 betaglycan fragment on TGF-β activity in MV 1 Lu cells. MV1 Lu assay was performed without (TGF-ß-) or with (TGF-ß+,0,1 ng/ml) added TGF-ß and various concentrations of bg3 betaglycan fragment or an unrelated fusion protein prepared in the same way as bg3 as a control. The bars show ± standard error of mean (n=4).

Figure 8 shows the specificity of the effect of the bg3 betaglycan fragment on TGF-ß induced growth suppression. MV 1 Lu assay was performed without or with TGF-ß (0.1 ng/ml), bg3 (20 μg/ml) and neutralizing anti-TGF-ß antibodies (20 μg/ml) as indicated in the figure. The bars show standard error of mean (n=4).

Figures 9A and 9B depict the nucleotide sequence [SEQ ID NO:1] and deduced amino acid sequence [SEQ ID NO:2] of a cDNA encoding rat betaglycan.

Figure 10 depicts the nucleotide sequence [SEQ ID NO:3] and deduced amino acid sequence [SEQ ID NO:4] of a cDNA encoding human betaglycan.

DETAILED DESCRIPTION OF THE INVENTION

This invention identifies, for the first time, a binding site in betaglycan for TGF-β. Results of experiments reported herein demonstrate that a polypeptide containing about one-fourth of the extracellular domain of betaglycan closest to the cell membrane has the TGF-β binding site. In particular, a polypeptide containing amino acids 543 to 769 of rat betaglycan [SEQ ID NO:2] contains this binding site. This polypeptide enhances the binding of TGF-β to the TGF-β type II receptor, enhances suppression of cell growth by TGF-β and competes with decorin-type proteoglycans for TGF-β binding.

This invention provides polypeptides of at least 155 amino acids that bind to TGF-β and that have a sequence consisting essentially of a sequence from a portion of a mammalian betaglycan within about one-third of the extracellular domain closest to the cell membrane. As used herein, the term "sequence" in reference to a polypeptide means the amino acid sequence of the polypeptide. A sequence consists essentially of a sequence of a mammalian betaglycan if it corresponds, or corresponds except for minor modifications, to a portion of a sequence of a mammalian betaglycan.

As used herein, "minor modifications" refers to simple substitutions, additions or deletions that do not eliminate the TGF-β binding capacity of the polypeptide or its ability to enhance TGF-β bioactivity. These modifications may be deliberate, as through site-directed mutagenesis, or may be accidental, such as through

mutation in hosts having DNA encoding these polypeptides. Simple substitutions include the substitution of an amino acid for another having a side chain off the alpha carbon of the same class, i.e. non-polar (hydrophobic), neutral, positively charged or negatively charged.

The polypeptides of this invention have a sequence from a mammalian betaglycan and, in particular, from human, rat or pig betaglycan. The nucleotide and amino acid sequences of rat betaglycan are given in Figures 9A-9B [SEQ ID NOS:1 and 2] and in Lopez-Casillas et al., *Cell*, 67:785-95 (1991) (incorporated herein by reference). The nucleotide and amino acid sequences of human betaglycan are given in Figure 10 [SEQ ID NOS:3 and 4] and in Moren et al., *Biochem. Biophys. Res. Comm.*, 189:356-362 (1992) (incorporated herein by reference). Moren et al. also provides the amino acid sequence of pig betaglycan.

About one-third of the extracellular domain of betaglycan closest to the cell membrane corresponds to about amino acid 501 to 769 of rat betaclycan [SEQ ID NO:2]. This corresponds to about amino acid 497 to 765 [SEQ ID NO:4] of human betaglycan. Other corresponding amino acids are apparent from a comparison of these two sequences.

Polypeptides of this invention include those wherein the portion of a mammalian betaglycan is about one-fifth of the extracellular domain of a mammalian betaglycan closest to the cell membrane. A one-hundred-fifty-five amino acid polypeptide having the sequence of amino acids 615-769 of SEQ ID NO:2 is one such polypeptide.

Preferably, polypeptides of this invention include those wherein the portion of a mammalian

betaglycan is about one-fourth of the extracellular domain of a mammalian betaglycan closest to the cell membrane. The two-hundred-twenty-seven amino acid polypeptide having the sequence of amino acids 543-769 of SEQ ID NO:2 is one such polypeptide.

Polypeptides of this invention include those having sequences consisting essentially of a portion of the sequence of rat betaglycan [SEQ ID NO:2]. According to one embodiment of the invention, the polypeptide has a sequence consisting essentially of at least amino acids 615 to 769 of SEQ ID NO:2 to at most amino acids 501 to 769 of SEQ ID NO:2.

More preferably, the polypeptide has a sequence consisting essentially of at least amino acids 543 to 769 of SEQ ID NO:2 to at most amino acids 501 to 769 of SEQ ID NO:2. All of these polypeptides include amino acids in a sequence shown to bind to TGF-β. They exclude polypeptides having the sequence that Lopez-Casillas et al., *Cell*, 73:1435-44 (1993) asserted bind to TGF-β.

This invention is also directed to soluble polypeptides having the formula: A-B-C, wherein A is an amino acid sequence that excludes sequences of more than 4 amino acids from amino acids 1 to 501 of SEQ ID NO 1; B is a sequence consisting essentially of at least amino acids 543 to 769 of SEQ ID NO:2 to at most amino acids 501 to 769 of SEQ ID NO:2; and C is an amino acid sequence. In one embodiment of the invention, neither A nor C have more than 100 amino acids. Soluble polypeptides are those lacking a transmembrane region, the hydrophobic region of the polypeptide that anchors it in the cell membrane.

This invention is also directed to soluble polypeptides having the formula: A-B-C, wherein A is an

amino acid sequence that excludes sequences of more than 4 amino acids from amino acids 1 to 543 of SEQ ID NO 1; B is at least 155 amino acids in a sequence consisting essentially of an amino acid sequence within amino acids 543 to 769 of SEQ ID NO:2; and C is an amino acid sequence. In one embodiment of this invention, neither A nor C have more than 100 amino acids.

The polypeptides of this invention can be produced by synthesis on an automated peptide synthesizer, according to the manufacturer's instructions. For example, MODEL 430A, Applied Biosystems, Foster City, California, USA, is a synthesizer useful for this purpose. The polypeptides of this invention can also be produced by the expression of a nucleic acid molecule that encodes the polypeptide. Methods for expressing the nucleic acids of this invention are described below.

This invention is also directed to isolated nucleic acids that encode any of the polypeptides of this invention. Nucleic acid molecules of this invention can have nucleotide sequences for portions of rat or human betaglycan derived from Figures 9A-9B [SEQ ID NO:2] or Figure 10 [SEQ ID NO:4]. For example, a nucleic acid having a sequence of at least nucleotides 1961 to 2641 and at most 1835 to 2641 of SEQ ID NO:1 encodes a polypeptide having a sequence of at least amino acids 543 to 769 of SEQ ID NO:2 to at most amino acids 501 to 769 of SEQ ID NO:2. Nucleic acid molecules of this invention include degenerate versions of sequences of mammalian genes.

This invention is further directed to expression vectors having an expression control sequence operatively linked to a nucleic acid of this invention. Expression vectors useful in this invention include

plasmids, cosmids, phage and the like. An expression control sequence is operatively linked to a nucleic acid molecule when it directs the transcription and translation of that molecule in an appropriate host cell. Expression vectors and their use are well known to the art. This invention is further directed to prokaryotic and eukaryotic cells transfected with an expression vector of this invention and capable of expressing the nucleic acid of this invention.

The nucleic acids of this invention can be produced by organic synthesis on a commercial nucleic acid synthesizer or through PCR on a nucleic acid encoding a mammalian betaglycan. Nucleic acid sequences encoding mammalian betaglycans, can be identified by probing cDNA libraries with probes derived from rat betaglycan [SEQ ID NO:1] or human betaglycan [SEQ ID NO:3] and by analyzing cDNA expression libraries with antibodies against betaglycan. Alternatively, betaglycan from these mammals can be isolated and partially sequenced, and the sequence can be used to make sets of degenerate nucleic acid probes for probing gene libraries. Other methods for identifying and isolating genes are also known.

The construction of expression vectors and the expression of genes in transfected cells involves the use of molecular cloning techniques also well known in the art. Sambrook et al., *Molecular Cloning -- A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, (1989) (incorporated herein by reference) provides many protocols in the art of molecular genetics.

The nucleic acids, expression vectors and cells of this invention are useful for producing the polypeptides of this invention. The nucleic acids of

this invention also find use as probes for detecting a nucleic acid having a sequence encoding betaglycan.

The polypeptides of this invention find use in methods of detecting TGF-β in a sample. Since levels of TGF-β are altered upon injury or other pathologies, the level of TGF-β is a useful sign of these conditions. The methods involve contacting the sample with a peptide of this invention and determining the amount of TGF-β bound to the polypeptide. According to one embodiment of this method, the well of a microtiter plate is coated with a polypeptide of this invention. The sample is added to the well and incubated under conditions to allow binding of TGF-β to the polypeptide. Unbound sample is removed. Then, the amount of TGF-β bound is determined by, for example, contacting the microtiter well with an anti-TGF-β antibody bound to a reporter group. Reporter groups useful in this invention include chemiluminescent labels, fluorescent labels, radioactive labels, enzyme labels and the like. Variations on this method will be apparent to any person skilled in the art.

The polypeptides of this invention find use in methods of isolating TGF-β from a sample by contacting the mixture with a peptide of this invention bound to a solid support to allow binding and isolating the TGF-β from the polypeptide. According to one embodiment of this method, a polypeptide of this invention is bound to an insoluble matrix and made into an affinity column. The sample is passed over the column under conditions to allow the binding of TGF-β to the polypeptide. Unbound material is washed out of the column. TGF-β is recovered by washing the column with a solution and under conditions that allow TGF-β to become unbound from the polypeptide. Variations on this method will be apparent to any person skilled in the art.

The polypeptides of this invention find use in methods of enhancing the binding of TGF-ß to type II receptor. These methods involve, for example, contacting a cell bearing a type II receptor with TGF-β and a polypeptide of this invention. This invention also provides methods of enhancing suppression of cell growth, particularly epithelial cell growth, by TGF-β. These methods involve, for example, contacting a cell with TGF-β and a polypeptide of this invention. Embodiments of these methods are described in the Example.

The induction of these cellular events finds use in therapeutic methods in which the events are preventative or ameliorative. In any such therapeutic method, the polypeptides of this invention normally will be administered in a pharmaceutical composition. In one embodiment of the invention, the pharmaceutical compositions comprise the polypeptides of this invention in a pharmaceutically acceptable carrier. The pharmaceutical composition can further comprise TGF-ß.

As used herein, the term "pharmaceutically acceptable carrier" encompasses any of the standard pharmaceutical carriers, such as a phosphate-buffered saline solution, water, and emulsions, such as an oil/water or water/oil emulsion, and various types of wetting agents. Suitable pharmaceutical carriers and their formulations are described in Martin, *Remington's Pharmaceutical Sciences*, 15th Ed. (Mack Publishing Co., Easton 1975). Such compositions will, in general, contain a therapeutically effective amount of the polypeptide together with a suitable amount of carrier so as to prepare the proper dosage form for proper administration to the subject.

This invention provides methods of treating a subject with a condition ameliorated by the enhanced

binding of TGF-β to type II receptor or by the enhanced suppression of cell growth by TGF-β by administering to the subject a therapeutically effective amount of a pharmaceutical composition having a polypeptide of this invention. The pharmaceutical composition can further include TGF-β.

As used herein, the term "therapeutically effective amount" is that amount necessary to alleviate the condition from which the subject suffers or prevent such a condition. As used herein, the term "subject" includes humans, other mammals or other vertebrates.

In the practice of the therapeutic methods of the present invention, an effective amount of a polypeptide of this invention, including derivatives or salts thereof, or a pharmaceutical composition containing the same, as described above, is administered via any of the usual and acceptable methods known in the art, either singly or in combination with other pharmaceutical agents.

In the practice of the therapeutic methods of the invention, the particular dosage of pharmaceutical composition to be administered to the subject will depend on a variety of considerations including the nature of the disease, the severity thereof, the schedule of administration, the age and physical characteristics of the subject, and so forth. Proper dosages may be established using clinical approaches familiar to the medicinal arts.

In the practice of the therapeutic methods of this invention, the pharmaceutical compositions can be administered via any of the usual and acceptable methods known in the art, for example orally, parenterally (e.g., intra-muscularly, intravenously, subcutaneously or

locally to other tissues) or by inhalation, and in the form of solid or liquid dosage including tablets, suspensions, and aerosols.

This invention also provides decoy betaglycan polypeptides. As used herein, a "decoy betaglycan polypeptide" is a polypeptide having a sequence corresponding to at least a portion of a mammalian betaglycan except for disabling modifications to the amino acid sequence. As used herein, "disabling modifications" refers to simple substitutions, additions or deletions that allow retention of the TGF-β binding capacity of the polypeptide but that eliminate its ability to enhance TGF-ß bioactivity, such as the binding of TGF-ß to a TGF-β receptor, such as the TGF-β type II receptor, or the suppression of cell growth by TGF-β. Disabling modifications also allow the decoy to suppress TGF-β bioactivity.

Decoy betaglycan polypeptides contain disabling modifications to betaglycan within the region known to enhance TGF-β bioactivity, that is, within about one-fourth of the extracellular domain of a mammalian betaglycan closest to the cell membrane and, in particular, to amino acids 543 to 769 of SEQ ID NO:2. These modifications can be introduced deliberately, as through site-directed mutagenesis.

Disabling modifications include, for example, the deletion of one or more amino acids, substitution of an amino acid for another having a different class of side chain off the alpha carbon, the elimination of a cysteine residue involved in disulfide bonding necessary for activity, the introduction of a proline or cysteine residue to alter the polypeptide's secondary structure and the like. Decoy betaglycan polypeptides can be identified by introducing likely disabling modifications

into the amino acid sequence and testing the resulting polypeptides for activity in any of the assays known to the art or described herein.

Decoy betaglycan polypeptides find use in detecting TGF-β in a sample and in isolating TGF-β from a sample. They also find use in methods of treating a subject with a condition ameliorated by the diminished binding of TGF-β to TGF-β receptor, by the inhibition of the suppression of cell growth by TGF-β, or the suppression of any other bioactivity of TGF-β. In particular, decoy betaglycan polypeptides find use in suppressing TGF-β-induced deposition of extracellular matrix at a site of tissue injury. Thus, they are useful for preventing or ameliorating glomerulonephritis, diabetic nephropathy, lung fibrosis, liver cirrhosis, intimal hyperplasia, cardiac cirrhosis after infarct, adult respiratory distress syndrome and other fibrosis-related pathologies. These methods involve administering to the subject a therapeutically effective amount of a pharmaceutical composition having a decoy betaglycan polypeptide of this invention.

This invention also provides anti-betaglycan-binding-site antibodies that eliminate the ability of betaglycan to enhance TGF-ß bioactivity. This includes the binding of TGF-ß to a TGF-β receptor, such as the TGF-β type II receptor, or the suppression of cell growth by TGF-β. Anti-betaglycan-binding-site antibodies find use in detecting betaglycan in a sample and in isolating betaglycan from a sample. They also find use in methods of treating a subject with a condition ameliorated by the diminished binding of TGF-β to TGF-β receptor, by the inhibition of the suppression of cell growth by TGF-β, or the suppression of any other bioactivity of TGF-β. In particular, anti-betaglycan binding site antibodies find use in suppressing TGF-β-induced deposition of

extracellular matrix at a site of tissue injury. Thus, they are useful for preventing or ameliorating glomerulonephritis, diabetic nephropathy, lung fibrosis, liver cirrhosis, intimal hyperplasia, cardiac cirrhosis after infarct, adult respiratory distress syndrome and other fibrosis-related pathologies. These methods involve administering to the subject a therapeutically effective amount of a pharmaceutical composition having an anti-betaglycan-binding-site antibody of this invention.

Anti-betaglycan-binding-site antibodies can be made by inoculating an animal with a polypeptide of this invention. For example, an aminal can be inoculated with a polypeptide having the sequence of amino acids 543 to 769 of rat betaglycan [SEQ ID NO:2].

An idiotype represents the specificity of an antibody for its ligand at the binding site. An anti-idiotypic antibody is an antibody directed against the idiotype of another antibody. They are made by immunizing an animal with the other antibody. Anti-idiotypic antibodies can be made that have the internal image of the antigen against which the other antibody is directed and that mimic the binding characteristics of the antigen.

Accordingly, this invention is also directed to anti-idiotypic antibodies that mimic the ability of the polypeptides of this invention to bind TGF-β. Such anti-idiotypic antibodies lack the biological function of betaglycan to enhance binding of TGF-β to the type II receptor or to enhance the suppression of cell growth by TGF-β. They also suppress other bioactivities of TGF-β. Anti-idiotypic antibodies of this invention can be made by inoculating an animal with an anti-betaglycan-binding-site antibody of this invention.

The anti-idiotypic antibodies of this invention find use in detecting TGF-β in a sample and in isolating TGF-β from a sample. They also find use in methods of treating a subject with a condition ameliorated by the diminished binding of TGF-β to a TGF-β receptor or by the inhibition of the suppression of cell growth by TGF-β. They also find use in methods of suppressing TGF-β-induced deposition of extracellular matrix at a site of tissue injury. Thus, they are useful for preventing or ameliorating conditions associated with this that are described above. The methods involve administering to a subject a therapeutically effective amount of a pharmaceutical composition having an anti-idiotypic antibody of this invention.

Methods of immunizing animals, isolating antibodies, and producing polyclonal or monoclonal antibodies are well known in the art and are described in, e.g., Harlow and Lane, Antibodies - A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, (1988). Kennedy et al., Scientific American, 255:48-56 (July 1986) and Carlsson and Glad, Bio/Technology, 7:567-73 (June 1989) describe anti-idiotypic antibodies. All of the above-identified texts are incorporated herein by reference.

The following Example is intended to illustrate but not limit the invention.

EXAMPLE

The pBluescript SK(+) vector and XL1 Blue host cells were purchased from STRATAGENE® (La Jolla, California). The expression vectors, pQE8, 10, 11 and M15 host cells and Ni-NTA-agarose came from QIAGEN® (Chatsworth, California), the buffer for PCR reactions from IDAHO TECHNOLOGY® (Idaho Falls, Idaho) and the Taq

polymerase from BOEHRINGER MANNHEIM®. GM 242 host cells (dam$^-$) were used in these experiments. Other dam$^-$ cell lines, for example, GM 48 or GM 163 cells (available from New England Biolabs, Beverly, Massachusetts) can also be used. The purification kits for PCR fragments and for the plasmid were from PROMEGA® (Madison, Connecticut). All other cloning reagents were from INTERNATIONAL BIOTECHNOLOGIES INC.® (La Jolla, California) and BIORAD® (Richmond, California).

TGF-ß1 is commercially available from Genzyme (Cambridge, Massachusetts). SDS-polyacrylamide gel electrophoresis (PAGE) pre-cast gels were from NOVEX® (San Diego, California). Hep G2 and Mv 1 Lu cells were from American Type Culture Collection (ATCC HB 8065 and ATCC CCL 64, respectively). Fetal calf serum was purchased from TISSUE CULTURE BIOLOGICALS® (Tulare, California) and L-glutamine, antibiotics and antimycotic agents from IRVINE SCIENTIFIC® (Santa Ana, California). DMEM (Dulbecco's modified Eagle's medium) was from GIBCO® (Gaithersburg, Maryland), carrier-free $Na^{125}I$ and ^3H-thymidine from NEW ENGLAND NUCLEAR® (Boston, Massachusetts), and IODO-GEN® from Pierce Chemical Co. (Rockford, Illinois). IMMULON 2 REMOVAWELL® strips came from Dynatech Laboratories Inc. (Chantilly, Virginia), and CENTRICON® micro-concentrator from Amicon (Danvers, Massachusetts). All chromatographic materials including pre-packed PD-10 columns were from PHARMACIA® (Uppsala, Sweden). Neutralizing chicken anti-TGF-ß antibody and normal chicken Ig were from R & D SYSTEMS® (Minneapolis, Minnesota).

The following segments from the DNA sequence of rat betaglycan [SEQ ID NO:1] were first amplified by PCR with appropriate tags for cloning purposes: nucleotides 404-1439 (A); nucleotides 1126-2268 (B); nucleotides 1712-3042 (C); and nucleotides 1961-2641 (D) (López-

Casillas et al., *Cell*, 67: 785-95 (1991)). PCR reactions were performed on cDNA derived from rat smooth muscle cells using a MODEL 1605® Air Thermo-Cycler (Idaho Technology, Idaho Falls, ID). The fragments were ligated into the TA cloning site of the pBluescript SK vector (Marchuk et al., *Nucl. Acids. Res.*, 19:1154 (1990)) and the recombinant vector was transformed into XL1 Blue. The plasmid carrying PCR product B was re-transformed into GM 242 to recover a demethylated *Bcl* I site. All transformations were carried out using the CELL-PORATOR® (BRL, Gaithersburg, MD) electroporation system, and positive colonies were selected by the addition of isopropyl-b-D-thiogalactopyranoside (IPTG) and X-gal. The sequences of the PCR-generated DNAs completely agreed with the published sequence for betaglycan (López-Casillas et al., 1991, *supra*; Wang et al., *Cell*, 67:792-805 (1991)).

From product A, fragment bg1 was excised by *Bam* HI digestion. Fragment bg2 was obtained from product B by *Bcl* I and *Sal* I digestion, and fragments bg3'N, bg3N and bg3C were prepared from product C by *Bam* HI, *Bam* HI and *Xmn* I, and *Bam* HI and *Bgl* II digestion, respectively. Product D was designated bg3. Fragments bg1, bg3, bg3'N and bg3N were ligated into the *Bam* HI site of the pQE8 expression vector. Fragments bg3C and bg4 were cloned into pQE10 and bg2 into pQE11. The vectors were transformed into M15 host cells by electroporation. Positive clones were selected by testing for the expression of the protein product.

The expression and purification of the bacterially expressed proteins were carried out according to instructions on page 16 of Qiagen's manual. Briefly, the expression was induced by 2 mM of IPTG and continued for 3 hr, after which the bacteria were lysed in 6 M guanidine-HCl, Tris buffer, pH 8. The supernatant was

loaded onto a column of Ni-NTA resin, and the column was eluted with 8 M urea in Tris buffer, pH 4.5, after several washing steps. The protein solution was neutralized with 1 M Tris buffer, pH 9, containing 8 M urea. The proteins were treated with dithiotreitol (50 °C for 30 min) and iodoacetamide (25 °C for 30 min) to reduce and alkylate cysteine residues. (Matsudaira, "A Practical Guide to Protein and Peptide Purification for Microsequencing," Academic Press, (1989)). The reaction mixture was loaded onto a PD-10 column equilibrated with 0.1 M ammonium bicarbonate containing 6 M urea. Elution of protein was monitored by protein assay (Bradford, Anal. Biochem., 72:248-54 (1976)), and the main fraction was concentrated with CENTRICON 10®.

Eighty-percent-confluent monolayers of cells maintained in DMEM containing 10% FCS were used for TGF-ß binding assays and affinity labeling experiments. The conditions used were essentially as described earlier, except that the labeling reaction, which is usually performed at +4 °C, was also carried out at 37 °C (Massagué, *Meth. Enz.*, 146:174-195 (1987)). Competitors were dissolved in 6 M urea at 150-fold molar excess of solution. Final samples were prepared by diluting the sample 150-fold into assay buffer containing the ^{125}I-TGF-ß1. The following buffers were used: 25 mM Hepes, pH 7.4, 125 mM NaCl, 5 mM $MgSO_4$, 5 mM KCl, 1 mM $CaCl_2$, 2 mg/ml bovine serum albumin (BSA) ("binding buffer"); 10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 % Triton X-100 ("solubilization buffer"); 10 mg/ml antipain, 10 mg/ml leupeptin, 2 mM benzamidine, 1 mM ethylenediamine-tetraacetic acid ("protease inhibitor cocktail," final concentrations). A fresh stock of disuccinimidyl suberate (10 mg/ml in dimethyl sulfoxide) was prepared for each experiment, and a 1:200 dilution in binding buffer without BSA was used in cross-linking for 15 minutes on ice. The reaction was quenched by 10 mM Tris-

HCl, pH 7.4, 150 mM NaCl and the samples were solubilized and analyzed.

Binding of ^{125}I-labeled proteins to immobilized TGF-ß1 was determined using a solid-phase binding assay (Mooradian et al., *J. Cell. Biochem.*, 41:189-200 ((1989)). Briefly, 96-well IMMULON 2® plates were coated with 1 mg/ml TGF-b1 (100 ml, 0.1 M carbonate buffer, pH 9.5) at 4 °C for 16 hr. The wells were blocked by incubating with the blocking buffer containing 2 % BSA, 0.005 % Tween 20, 0.02 % NaN_3 for 2 hr at 37 °C. An equal volume of varying concentrations of competitors, diluted in 6 M urea, was mixed in the assay buffer (same as the blocking buffer) containing the labeled proteins. Since the binding was found to be sensitive to urea, the final urea concentration was always kept under 100 mM. Duplicate 100 ml samples were incubated in the wells for 5 hr at 37 °C. After four washings with the same buffer, the bound radioactivity was measured in a Beckman Gamma-4000 counter.

Data from the binding assays were analyzed using a *LIGAND* computer program (Munson and Rodbard, *Anal. Biochem.*, 107:220-239 (1980)) on an Apple Macintosh II computer.

The mink lung epithelial cell growth inhibition assay was done as described (Danielpour et al., *J. Cell. Physiol.*, 138:79-86 (1989)). Briefly, cells were grown in Dulbecco's medium containing 10 % fetal calf serum, 10 mM L-glutamine, 100 IU/ml penicillin and 100 mg/ml streptomycin. For the experiments, the cells were plated on 96-well plates (20,000 cells/well) in the same medium. After one day the medium was replaced by Dulbecco's medium containing 1 % fetal calf serum with glutamine and antibiotics. The experiments were started the next day by adding the controls and effectors to the cells in the

above medium. The fusion proteins were dialyzed against assay medium. After 24 hours of incubation the cells were pulsed with ^{3}H-thymidine for 3 hours. The cells were fixed with methanol and extracted 3 times with 10 % trichloracetic acid, solubilized with 1 % SDS-0.3 N NaOH and the radioactivity was counted.

SDS-PAGE was performed according to Laemmli (*Nature*, 227:680-685 (1970)) using vertical precast gels. For autoradiography after electrophoresis the gels were fixed in 10 % isopropanol, 10 % acetic acid for 15 minutes and dried. Kodak X-OMAT/AR® film with an enhancing screen was used for autoradiography.

Each of the pairs of PCR primers from the published rat betaglycan sequence generated a DNA fragment of the expected size. Shorter DNA fragments were prepared from the primary PCR products as shown in Figure 1A and ligated into a vector that expresses the cloned protein as a fusion protein with a cassette of six histidines. The expression system yielded 0.5 mg to 1.0 mg each of the betaglycan protein fragments from a 250 ml culture. The bg1 fragment contained a second band that may represent a fragment of the fusion protein; the other fusion protein products were essentially homogeneous. (Figure 1B.)

Betaglycan fragments bg1, bg2 and bg3 were tested for their ability to inhibit the binding of ^{125}I-TGF-b1 to Hep G2 cells; in these cells most of the specific binding of TGF-ß is to betaglycan. Hep G2 cells on 24-well dishes were incubated with ^{125}I-TGF-ß1 (100 pM) and with indicated concentrations of betaglycan fragments that had been expressed as fusion protein in bacteria. The bound TGF-ß was measured. Ten percent of the total labeled protein bound to the cells in the absence of any

competitor, and more than 90 % of this binding was displaced by the addition of 40 nM of unlabeled TGF-ß1.

Among the betaglycan fragments, only bg3 inhibited the binding of ^{125}I-TGF-ß to Hep G2 and the inhibition was concentration dependent. (Figure 2A.) The inhibitory activity of the bg3 fragment was lost upon further fragmentation. (Figure 2A.) Combinations of the sub-fragments also showed little or no effect. (Figure 2B.) There was some apparent enhancement of TGF-ß binding by the bg1 and bg2 fragments.

In affinity cross-linking of ^{125}I-TGF-ß to the cell surface receptors of MV 1 Lu cells, the bg3 fragment enhanced the binding of TGF-β to the bands representing the type II receptor and to betaglycan itself at low concentrations (up to 50 nM) of the fragment. (Figure 3.) Higher concentrations competed for the binding of TGF-ß to these binding sites. The cross-linking of TGF-ß to RI was unaffected by bg3 concentrations up to 500 nM, but higher concentrations appeared to be inhibitory.

The availability of the TGF-ß-binding betaglycan fragment were used to study various parameters of its binding to TGF-ß. A solid-phase binding assay in which ^{125}I-labeled betaglycan fragment binds to immobilized TGF-ß showed that saturation of binding was approached at the coating concentration of 10 mg/ml of TGF-ß1. (Figure 4A.) The binding was inhibited by 1 mM of unlabeled bg3 fragment. The binding of the bg3 fragment to TGF-ß reached a maximum (about 25 % of total added) after 5 hours of incubation when 1 mg/ml of TGF-ß1 was used for the coating. (Figure 4B.)

Homologous competition for the binding of ^{125}I-bg3 to immobilized TGF-ß1 by various concentrations of unlabeled bg3 was performed. (Figure 5A.) Scatchard

analysis of these results showed an excellent theoretical fit with a two-site binding model giving two dissociation constants, K_d = 3.9 nM and K_d = 145 nM. (Figure 5B.)

The core proteins of extracellular matrix proteoglycans decorin, biglycan and fibromodulin bind TGF-ß. Solid-phase binding assay, used to compare the TGF-ß binding characteristics of these proteoglycans and betaglycan, showed that fusion proteins representing the core proteins of decorin, biglycan and fibromodulin each inhibited the binding of the bg3 betaglycan fragment to TGF-β, whereas the fusion partner, maltose-binding protein, showed no significant effect. (Figure 6.) The 50 % inhibitory concentrations for the extracellular matrix proteoglycans were similar or slightly higher than for the betaglycan fragment. In a reverse experiment, the binding of labeled biglycan and fibromodulin core proteins to immobilized TGF-β1 was inhibited by unlabeled betaglycan fragment.

The bg3 betaglycan fragment, when administered together with a sub-maximally effective concentration of TGF-β1, enhanced the activity of TGF-ß1 in a concentration-dependent fashion. (Figure 7.) The fragment alone had no effect on DNA synthesis of MV 1 Lu cells. An unrelated fusion protein used as a negative control had no effect. The TGF-ß-promoting effect of bg3 could be blocked by adding neutralizing anti-TGF-ß antibodies into the assay. (Figure 8.) Non-immune IgG, which was used as a control, did not have this effect.

Little activity was found in the subfragments of the 30 kilodalton bg3 fragment in the assay that measures the ability of the fragments to inhibit the binding of TGF-β to cell surface receptors. This suggests that the TGF-β binding site in betaglycan may be an extended structure or that the binding site had lost

its proper conformation in the subfragments, resulting in a decrease of affinity.

Assays described above showed that betaglycan and the decorin-type small extracellular matrix proteoglycans interfered with one another's binding to TGF-β. This result suggests that the decorin-type proteoglycans bind to the same or overlapping site in TGF-β as betaglycan. That this is the case is also indicated by the finding that the decorin-type proteoglycans compete with the binding of TGF-β to betaglycan in receptor affinity cross-linking experiments.

Even though endoglin and betaglycan share sequence similarities and bind to TGF-β in a similar manner, the region in betaglycan where we have localized the TGF-β binding site shows no apparent similarity with any part of endoglin. However, the decorin-type proteoglycans also show no sequence similarity with betaglycan or endoglin. Moreover, the binding specificity of endoglin is different from that of betaglycan in that endoglin does not bind TGF-β2, whereas betaglycan does. It may be that the sequence similarities in betaglycan and endoglin relate to shared functions of these molecules other than the TGF-β binding.

One secondary function of betaglycan and possibly also endoglin, is the modulation of TGF-β binding to the signal transduction receptors. We found that the active betaglycan fragment could increase the binding of TGF-β to the type II receptor in cell surface affinity labeling and that the fragment enhanced the bioactivity of TGF-β when added to cell cultures as a soluble protein. The enhancement of the type II receptor binding is in agreement with the results of López-

Casillas et al. (1991), *supra*, who found that expression of recombinant betaglycan in a cell line that originally had little of it increased the TGF-β binding activity of the type II receptor.

It is noteworthy that our affinity labeling experiments showed inhibition of TGF-β binding to the type II receptor at high concentrations of the betaglycan fragment, even though the bioassay showed increased activity at those same concentrations. The explanation for these results may be that the two assays are done under different conditions; we were able to eliminate some of the differences in the test conditions but not others. We found that when the affinity labeling was performed at +4°, which is customary, the betaglycan fragment inhibited the affinity labeling of the type II receptor at all concentrations of the fragment (result not shown). However, when the temperature was raised to the 37° bioassay temperature, low and moderate concentrations of the fragment, now in agreement with the bioassay results, enhanced type II receptor labeling. The remaining major difference in the two assays, the time of incubation (10 minutes versus 24 hours), may explain the discrepancy of the results at the high fragment concentrations.

Although the invention has been described with reference to the presently-preferred embodiments, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the claims that follow.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: La Jolla Cancer Research Foundation

(ii) TITLE OF INVENTION: BETAGLYCAN POLYPEPTIDES HAVING TGF-BETA BINDING ACTIVITY

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Campbell and Flores
(B) STREET: 4370 La Jolla Village Drive, Suite 700
(C) CITY: San Diego
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 92122

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 14-OCT-1994
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Campbell, Cathryn A.
(B) REGISTRATION NUMBER: 31,815
(C) REFERENCE/DOCKET NUMBER: FP-LA 1185

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (619) 535-9001
(B) TELEFAX: (619) 535-8949

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3237 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 241..2799

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAGGAGGTGA AAGTCCCCGG CGGTCCGGAT GGCGCAGTTG CACTGCGCTG CTGAGCTCGC 60

GGCCGCCTGC GCACACTGGG GGGACTCGCT TCGGCTAGTA ACTCCTCCAC CTCGCGGCGG 120

ACGACCGGTC CTGGACACGC TGCCTGCGAG GCAAGTTGAA CAGTGCAGAG AAGGATCTTA 180

AAGCTACACC CGACTTGCCA CGATTGCCTT CAATCTGAAG AACCAAAGGC TGTTGGAGAG 240

```
ATG GCA GTG ACA TCC CAC CAC ATG ATC CCG GTG ATG GTT GTC CTG ATG     288
Met Ala Val Thr Ser His His Met Ile Pro Val Met Val Val Leu Met
1               5                   10                  15

AGC GCC TGC CTG GCC ACC GCC GGT CCA GAG CCC AGC ACC CGG TGT GAA     336
Ser Ala Cys Leu Ala Thr Ala Gly Pro Glu Pro Ser Thr Arg Cys Glu
            20                  25                  30

CTG TCA CCA ATC AAC GCC TCT CAC CCA GTC CAG GCC TTG ATG GAG AGC     384
Leu Ser Pro Ile Asn Ala Ser His Pro Val Gln Ala Leu Met Glu Ser
        35                  40                  45

TTC ACC GTT CTG TCT GGC TGT GCC AGC AGA GGC ACC ACC GGG CTG CCA     432
Phe Thr Val Leu Ser Gly Cys Ala Ser Arg Gly Thr Thr Gly Leu Pro
    50                  55                  60

AGG GAG GTC CAT GTC CTA AAC CTC CGA AGT ACA GAT CAG GGA CCA GGC     480
Arg Glu Val His Val Leu Asn Leu Arg Ser Thr Asp Gln Gly Pro Gly
65                  70                  75                  80

CAG CGG CAG AGA GAG GTT ACC CTG CAC CTG AAC CCC ATT GCC TCG GTG     528
Gln Arg Gln Arg Glu Val Thr Leu His Leu Asn Pro Ile Ala Ser Val
                85                  90                  95

CAC ACT CAC CAC AAA CCT ATC GTG TTC CTG CTC AAC TCC CCC CAG CCC     576
His Thr His His Lys Pro Ile Val Phe Leu Leu Asn Ser Pro Gln Pro
            100                 105                 110

CTG GTG TGG CAT CTG AAG ACG GAG AGA CTG GCC GCT GGT GTC CCC AGA     624
Leu Val Trp His Leu Lys Thr Glu Arg Leu Ala Ala Gly Val Pro Arg
        115                 120                 125

CTC TTC CTG GTT TCG GAG GGT TCT GTG GTC CAG TTT CCA TCA GGA AAC     672
Leu Phe Leu Val Ser Glu Gly Ser Val Val Gln Phe Pro Ser Gly Asn
    130                 135                 140

TTC TCC TTG ACA GCA GAA ACA GAG GAA AGG AAT TTC CCT CAA GAA AAT     720
Phe Ser Leu Thr Ala Glu Thr Glu Glu Arg Asn Phe Pro Gln Glu Asn
145                 150                 155                 160

GAA CAT CTC GTG CGC TGG GCC CAA AAG GAA TAT GGA GCA GTG ACT TCG     768
Glu His Leu Val Arg Trp Ala Gln Lys Glu Tyr Gly Ala Val Thr Ser
                165                 170                 175

TTC ACT GAA CTC AAG ATA GCA AGA AAC ATC TAT ATT AAA GTG GGA GAA     816
Phe Thr Glu Leu Lys Ile Ala Arg Asn Ile Tyr Ile Lys Val Gly Glu
            180                 185                 190

GAT CAA GTG TTT CCT CCT ACG TGT AAC ATA GGG AAG AAT TTC CTC TCA     864
Asp Gln Val Phe Pro Pro Thr Cys Asn Ile Gly Lys Asn Phe Leu Ser
        195                 200                 205

CTC AAT TAC CTT GCC GAG TAC CTT CAA CCC AAA GCC GCC GAA GGT TGT     912
Leu Asn Tyr Leu Ala Glu Tyr Leu Gln Pro Lys Ala Ala Glu Gly Cys
    210                 215                 220

GTC CTG CCC AGT CAG CCC CAT GAA AAG GAA GTA CAC ATC ATC GAG TTA     960
Val Leu Pro Ser Gln Pro His Glu Lys Glu Val His Ile Ile Glu Leu
225                 230                 235                 240

ATT ACC CCC AGC TCG AAC CCT TAC AGC GCT TTC CAG GTG GAT ATA ATA    1008
Ile Thr Pro Ser Ser Asn Pro Tyr Ser Ala Phe Gln Val Asp Ile Ile
                245                 250                 255

GTT GAC ATA CGA CCT GCT CAA GAG GAT CCC GAG GTG GTC AAA AAC CTT    1056
Val Asp Ile Arg Pro Ala Gln Glu Asp Pro Glu Val Val Lys Asn Leu
            260                 265                 270
```

```
GTC CTG ATC TTG AAG TGC AAA AAG TCT GTC AAC TGG GTG ATC AAG TCT      1104
Val Leu Ile Leu Lys Cys Lys Lys Ser Val Asn Trp Val Ile Lys Ser
        275                 280                 285

TTT GAC GTC AAG GGA AAC TTG AAA GTC ATT GCT CCC AAC AGT ATC GGC      1152
Phe Asp Val Lys Gly Asn Leu Lys Val Ile Ala Pro Asn Ser Ile Gly
    290                 295                 300

TTT GGA AAA GAG AGT GAA CGA TCC ATG ACA ATG ACC AAA TTG GTA AGA      1200
Phe Gly Lys Glu Ser Glu Arg Ser Met Thr Met Thr Lys Leu Val Arg
305                 310                 315                 320

GAT GAC ATC CCT TCC ACC CAA GAG AAT CTG ATG AAG TGG GCA CTG GAC      1248
Asp Asp Ile Pro Ser Thr Gln Glu Asn Leu Met Lys Trp Ala Leu Asp
                325                 330                 335

AAT GGC TAC AGG CCA GTG ACG TCA TAC ACA ATG GCT CCC GTG GCT AAT      1296
Asn Gly Tyr Arg Pro Val Thr Ser Tyr Thr Met Ala Pro Val Ala Asn
            340                 345                 350

AGA TTT CAT CTT CGG CTT GAG AAC AAC GAG GAG ATG AGA GAT GAG GAA      1344
Arg Phe His Leu Arg Leu Glu Asn Asn Glu Glu Met Arg Asp Glu Glu
        355                 360                 365

GTC CAC ACC ATT CCT CCT GAG CTT CGT ATC CTG CTG GAC CCT GAC CAC      1392
Val His Thr Ile Pro Pro Glu Leu Arg Ile Leu Leu Asp Pro Asp His
    370                 375                 380

CCG CCC GCC CTG GAC AAC CCA CTC TTC CCA GGA GAG GGA AGC CCA AAT      1440
Pro Pro Ala Leu Asp Asn Pro Leu Phe Pro Gly Glu Gly Ser Pro Asn
385                 390                 395                 400

GGT GGT CTC CCC TTT CCA TTC CCG GAT ATC CCC AGG AGA GGC TGG AAG      1488
Gly Gly Leu Pro Phe Pro Phe Pro Asp Ile Pro Arg Arg Gly Trp Lys
                405                 410                 415

GAG GGC GAA GAT AGG ATC CCC CGG CCA AAG CAG CCC ATC GTT CCC AGT      1536
Glu Gly Glu Asp Arg Ile Pro Arg Pro Lys Gln Pro Ile Val Pro Ser
            420                 425                 430

GTT CAA CTG CTT CCT GAC CAC CGA GAA CCA GAA GAA GTG CAA GGG GGC      1584
Val Gln Leu Leu Pro Asp His Arg Glu Pro Glu Glu Val Gln Gly Gly
        435                 440                 445

GTG GAC ATC GCC CTG TCA GTC AAA TGT GAC CAT GAA AAG ATG GTC GTG      1632
Val Asp Ile Ala Leu Ser Val Lys Cys Asp His Glu Lys Met Val Val
    450                 455                 460

GCT GTA GAC AAA GAC TCT TTC CAG ACC AAT GGC TAC TCA GGG ATG GAG      1680
Ala Val Asp Lys Asp Ser Phe Gln Thr Asn Gly Tyr Ser Gly Met Glu
465                 470                 475                 480

CTC ACC CTG TTG GAT CCT TCG TGT AAA GCC AAA ATG AAT GGT ACT CAC      1728
Leu Thr Leu Leu Asp Pro Ser Cys Lys Ala Lys Met Asn Gly Thr His
                485                 490                 495

TTT GTT CTC GAG TCT CCC CTG AAT GGC TGT GGT ACT CGA CAT CGG AGG      1776
Phe Val Leu Glu Ser Pro Leu Asn Gly Cys Gly Thr Arg His Arg Arg
            500                 505                 510

TCG ACC CCG GAT GGT GTG GTT TAC TAT AAC TCT ATT GTG GTG CAG GCT      1824
Ser Thr Pro Asp Gly Val Val Tyr Tyr Asn Ser Ile Val Val Gln Ala
        515                 520                 525

CCG TCC CCT GGG GAT AGC AGT GGC TGG CCT GAT GGC TAT GAA GAC TTG      1872
Pro Ser Pro Gly Asp Ser Ser Gly Trp Pro Asp Gly Tyr Glu Asp Leu
    530                 535                 540
```

```
GAG TCA GGC GAT AAT GGA TTT CCT GGA GAC GGG GAT GAA GGA GAA ACT      1920
Glu Ser Gly Asp Asn Gly Phe Pro Gly Asp Gly Asp Glu Gly Glu Thr
545                 550                 555                 560

GCC CCC CTG AGC CGA GCT GGA GTG GTG GTG TTT AAC TGC AGC TTG CGG      1968
Ala Pro Leu Ser Arg Ala Gly Val Val Val Phe Asn Cys Ser Leu Arg
                565                 570                 575

CAG CTG AGG AAT CCC AGT GGC TTC CAG GGC CAG CTC GAT GGA AAT GCT      2016
Gln Leu Arg Asn Pro Ser Gly Phe Gln Gly Gln Leu Asp Gly Asn Ala
            580                 585                 590

ACC TTC AAC ATG GAG CTG TAT AAC ACA GAC CTC TTT CTG GTG CCC TCC      2064
Thr Phe Asn Met Glu Leu Tyr Asn Thr Asp Leu Phe Leu Val Pro Ser
        595                 600                 605

CCA GGG GTC TTC TCT GTG GCA GAG AAC GAG CAT GTT TAT GTT GAG GTG      2112
Pro Gly Val Phe Ser Val Ala Glu Asn Glu His Val Tyr Val Glu Val
    610                 615                 620

TCT GTC ACC AAG GCT GAC CAA GAT CTG GGA TTC GCC ATC CAA ACC TGC      2160
Ser Val Thr Lys Ala Asp Gln Asp Leu Gly Phe Ala Ile Gln Thr Cys
625                 630                 635                 640

TTT CTC TCT CCA TAC TCC AAC CCA GAC AGA ATG TCT GAT TAC ACC ATC      2208
Phe Leu Ser Pro Tyr Ser Asn Pro Asp Arg Met Ser Asp Tyr Thr Ile
                645                 650                 655

ATC GAG AAC ATC TGT CCG AAA GAC GAC TCT GTG AAG TTC TAC AGC TCC      2256
Ile Glu Asn Ile Cys Pro Lys Asp Asp Ser Val Lys Phe Tyr Ser Ser
            660                 665                 670

AAG AGA GTG CAC TTT CCC ATC CCG CAT GCT GAG GTG GAC AAG AAG CGC      2304
Lys Arg Val His Phe Pro Ile Pro His Ala Glu Val Asp Lys Lys Arg
        675                 680                 685

TTC AGC TTC CTG TTC AAG TCT GTG TTC AAC ACC TCC CTG CTC TTC CTG      2352
Phe Ser Phe Leu Phe Lys Ser Val Phe Asn Thr Ser Leu Leu Phe Leu
    690                 695                 700

CAC TGC GAG TTG ACT CTG TGC TCC AGG AAG AAG GGC TCC CTG AAG CTG      2400
His Cys Glu Leu Thr Leu Cys Ser Arg Lys Lys Gly Ser Leu Lys Leu
705                 710                 715                 720

CCG AGG TGT GTC ACT CCT GAC GAC GCC TGC ACT TCT CTC GAT GCC ACC      2448
Pro Arg Cys Val Thr Pro Asp Asp Ala Cys Thr Ser Leu Asp Ala Thr
                725                 730                 735

ATG ATC TGG ACC ATG ATG CAG AAT AAG AAG ACA TTC ACC AAG CCC CTG      2496
Met Ile Trp Thr Met Met Gln Asn Lys Lys Thr Phe Thr Lys Pro Leu
            740                 745                 750

GCT GTG GTC CTC CAG GTA GAC TAT AAA GAA AAT GTT CCC AGC ACT AAG      2544
Ala Val Val Leu Gln Val Asp Tyr Lys Glu Asn Val Pro Ser Thr Lys
        755                 760                 765

GAT TCC AGT CCA ATT CCT CCT CCT CCT CCA CAG ATT TTC CAT GGC CTG      2592
Asp Ser Ser Pro Ile Pro Pro Pro Pro Pro Gln Ile Phe His Gly Leu
    770                 775                 780

GAC ACG CTC ACC GTG ATG GGC ATT GCA TTT GCA GCA TTT GTG ATC GGA      2640
Asp Thr Leu Thr Val Met Gly Ile Ala Phe Ala Ala Phe Val Ile Gly
785                 790                 795                 800

GCG CTC CTG ACG GGG GCC TTG TGG TAC ATC TAC TCC CAC ACA GGG GAG      2688
Ala Leu Leu Thr Gly Ala Leu Trp Tyr Ile Tyr Ser His Thr Gly Glu
                805                 810                 815
```

ACA GCA CGA AGG CAG CAA GTC CCT ACC TCG CCG CCA GCC TCG GAG AAC 2736
Thr Ala Arg Arg Gln Gln Val Pro Thr Ser Pro Pro Ala Ser Glu Asn
820 825 830

AGC AGC GCG GCC CAC AGC ATC GGC AGC ACT CAG AGT ACC CCC TGC TCT 2784
Ser Ser Ala Ala His Ser Ile Gly Ser Thr Gln Ser Thr Pro Cys Ser
835 840 845

AGC AGC AGC ACA GCC TAGGTGGACA GACAGACGCC CGCCCACCGC AGCCAGGGCA 2839
Ser Ser Ser Thr Ala
850

GGGCCCGATG CCAGTGCTGC GTGTCCACAG TCAGAAGTCT TGATCTGGGC TCCCTGTAAA 2899

GAAAGAGTGA ATTTCAGTAT ACAGACAGCC AGTTCTACCC ACCCCTTACC ACGGCCCACA 2959

TAAATGTGAC CCTGGGCATC TGTCACACGA AAGCTAAGCT GGTGGCCTTC CCCACCAGCC 3019

CCTCGCAGGA TGGGGGTTTC AATGTGAAAC ATCTGCCAGT TTTGTTTTGT TTTTTTAATG 3079

CTGCTTTGTC CAGGTGTCCA AACATCCATC ATTTGGGGTG GTCTGTTTTA CAGAGTAAAG 3139

GAGGCGGTGA AGGGACGTCA GCTAGTGTGT AGAGCCAAGG GGAGACAGCT AGGATTCTCG 3199

CCTAGCTGAA CCAAGGTGTA AAATAGAAGA CACGCTCC 3237

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 853 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Val Thr Ser His His Met Ile Pro Val Met Val Val Leu Met
1 5 10 15

Ser Ala Cys Leu Ala Thr Ala Gly Pro Glu Pro Ser Thr Arg Cys Glu
20 25 30

Leu Ser Pro Ile Asn Ala Ser His Pro Val Gln Ala Leu Met Glu Ser
35 40 45

Phe Thr Val Leu Ser Gly Cys Ala Ser Arg Gly Thr Thr Gly Leu Pro
50 55 60

Arg Glu Val His Val Leu Asn Leu Arg Ser Thr Asp Gln Gly Pro Gly
65 70 75 80

Gln Arg Gln Arg Glu Val Thr Leu His Leu Asn Pro Ile Ala Ser Val
85 90 95

His Thr His His Lys Pro Ile Val Phe Leu Leu Asn Ser Pro Gln Pro
100 105 110

Leu Val Trp His Leu Lys Thr Glu Arg Leu Ala Ala Gly Val Pro Arg
115 120 125

Leu Phe Leu Val Ser Glu Gly Ser Val Val Gln Phe Pro Ser Gly Asn
130 135 140

Phe Ser Leu Thr Ala Glu Thr Glu Glu Arg Asn Phe Pro Gln Glu Asn
145 150 155 160

```
Glu His Leu Val Arg Trp Ala Gln Lys Glu Tyr Gly Ala Val Thr Ser
                165                 170                 175

Phe Thr Glu Leu Lys Ile Ala Arg Asn Ile Tyr Ile Lys Val Gly Glu
            180                 185                 190

Asp Gln Val Phe Pro Pro Thr Cys Asn Ile Gly Lys Asn Phe Leu Ser
        195                 200                 205

Leu Asn Tyr Leu Ala Glu Tyr Leu Gln Pro Lys Ala Ala Glu Gly Cys
    210                 215                 220

Val Leu Pro Ser Gln Pro His Glu Lys Glu Val His Ile Ile Glu Leu
225                 230                 235                 240

Ile Thr Pro Ser Ser Asn Pro Tyr Ser Ala Phe Gln Val Asp Ile Ile
                245                 250                 255

Val Asp Ile Arg Pro Ala Gln Glu Asp Pro Glu Val Val Lys Asn Leu
            260                 265                 270

Val Leu Ile Leu Lys Cys Lys Lys Ser Val Asn Trp Val Ile Lys Ser
        275                 280                 285

Phe Asp Val Lys Gly Asn Leu Lys Val Ile Ala Pro Asn Ser Ile Gly
    290                 295                 300

Phe Gly Lys Glu Ser Glu Arg Ser Met Thr Met Thr Lys Leu Val Arg
305                 310                 315                 320

Asp Asp Ile Pro Ser Thr Gln Glu Asn Leu Met Lys Trp Ala Leu Asp
                325                 330                 335

Asn Gly Tyr Arg Pro Val Thr Ser Tyr Thr Met Ala Pro Val Ala Asn
            340                 345                 350

Arg Phe His Leu Arg Leu Glu Asn Asn Glu Glu Met Arg Asp Glu Glu
        355                 360                 365

Val His Thr Ile Pro Pro Glu Leu Arg Ile Leu Leu Asp Pro Asp His
    370                 375                 380

Pro Pro Ala Leu Asp Asn Pro Leu Phe Pro Gly Glu Gly Ser Pro Asn
385                 390                 395                 400

Gly Gly Leu Pro Phe Pro Phe Pro Asp Ile Pro Arg Arg Gly Trp Lys
                405                 410                 415

Glu Gly Glu Asp Arg Ile Pro Arg Pro Lys Gln Pro Ile Val Pro Ser
            420                 425                 430

Val Gln Leu Leu Pro Asp His Arg Glu Pro Glu Glu Val Gln Gly Gly
        435                 440                 445

Val Asp Ile Ala Leu Ser Val Lys Cys Asp His Glu Lys Met Val Val
    450                 455                 460

Ala Val Asp Lys Asp Ser Phe Gln Thr Asn Gly Tyr Ser Gly Met Glu
465                 470                 475                 480

Leu Thr Leu Leu Asp Pro Ser Cys Lys Ala Lys Met Asn Gly Thr His
                485                 490                 495

Phe Val Leu Glu Ser Pro Leu Asn Gly Cys Gly Thr Arg His Arg Arg
            500                 505                 510
```

```
Ser Thr Pro Asp Gly Val Val Tyr Tyr Asn Ser Ile Val Val Gln Ala
        515                 520                 525

Pro Ser Pro Gly Asp Ser Ser Gly Trp Pro Asp Gly Tyr Glu Asp Leu
    530                 535                 540

Glu Ser Gly Asp Asn Gly Phe Pro Gly Asp Gly Asp Glu Gly Glu Thr
545                 550                 555                 560

Ala Pro Leu Ser Arg Ala Gly Val Val Val Phe Asn Cys Ser Leu Arg
                565                 570                 575

Gln Leu Arg Asn Pro Ser Gly Phe Gln Gly Gln Leu Asp Gly Asn Ala
            580                 585                 590

Thr Phe Asn Met Glu Leu Tyr Asn Thr Asp Leu Phe Leu Val Pro Ser
        595                 600                 605

Pro Gly Val Phe Ser Val Ala Glu Asn Glu His Val Tyr Val Glu Val
    610                 615                 620

Ser Val Thr Lys Ala Asp Gln Asp Leu Gly Phe Ala Ile Gln Thr Cys
625                 630                 635                 640

Phe Leu Ser Pro Tyr Ser Asn Pro Asp Arg Met Ser Asp Tyr Thr Ile
                645                 650                 655

Ile Glu Asn Ile Cys Pro Lys Asp Asp Ser Val Lys Phe Tyr Ser Ser
            660                 665                 670

Lys Arg Val His Phe Pro Ile Pro His Ala Glu Val Asp Lys Lys Arg
        675                 680                 685

Phe Ser Phe Leu Phe Lys Ser Val Phe Asn Thr Ser Leu Leu Phe Leu
    690                 695                 700

His Cys Glu Leu Thr Leu Cys Ser Arg Lys Lys Gly Ser Leu Lys Leu
705                 710                 715                 720

Pro Arg Cys Val Thr Pro Asp Asp Ala Cys Thr Ser Leu Asp Ala Thr
                725                 730                 735

Met Ile Trp Thr Met Met Gln Asn Lys Lys Thr Phe Thr Lys Pro Leu
            740                 745                 750

Ala Val Val Leu Gln Val Asp Tyr Lys Glu Asn Val Pro Ser Thr Lys
        755                 760                 765

Asp Ser Ser Pro Ile Pro Pro Pro Pro Pro Gln Ile Phe His Gly Leu
    770                 775                 780

Asp Thr Leu Thr Val Met Gly Ile Ala Phe Ala Ala Phe Val Ile Gly
785                 790                 795                 800

Ala Leu Leu Thr Gly Ala Leu Trp Tyr Ile Tyr Ser His Thr Gly Glu
                805                 810                 815

Thr Ala Arg Arg Gln Gln Val Pro Thr Ser Pro Pro Ala Ser Glu Asn
            820                 825                 830

Ser Ser Ala Ala His Ser Ile Gly Ser Thr Gln Ser Thr Pro Cys Ser
        835                 840                 845

Ser Ser Ser Thr Ala
    850
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4213 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 622..3168

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GCGGACGCGA TCGCAACCTC CGCTTCCCGG GTTCAAGGGA TTCTCCTGCC TCAGCCTCCT        60

GAGTAGCTGG GACTACAGGC GCACCACTAC ACCCGGCCAA TTTTTGTATT TATTTATTTT       120

ATTTTATTTT ATTTTTTAGT AGAGACGGGG TTTACCATGT TGGCCAGAAC AGTCTCGATC       180

TCCTAACCTC GTGATCCAAC GCCTCGGCTC CCAAAGTGTT GAGTTACAGG CGTGAGCCAC       240

CGCGCCCGTC CAGAATTATC TTTAAAATTT ATTTCTTTAA GATTTGTAGC TACTAAGAAA       300

GAAAGGAGCT TTTTTTCCTT GGGCCTTCAA ACTGAAAGAA CCGCATGAGC CTGACGGCGC       360

ATGGTCTTAA CATCAGGCTG TGCAGGAAGA AGCTATCTGC AGATGGATGC CAGCACACAC       420

AAGGAAGCAG AGCTCTGGCA ACATTGAGTC AAAGCAAGGA CACAACATCA GAGGGACGGC       480

AGAGAATCCT TGTGTGTAGT CTTTGGTGGC AGTTTGAAAA TTGCAAGGAG GGACTTTAAG       540

ACTACTTCTG ATTTGCAAAG ATGGTCTGTG CTCCGAGCAG GCTAAAGTGA CTGGACGAGA       600

CGCACTGTTG GAGAAATAAA A ATG ACT TCC CAT TAT GTG ATT GCC ATC TTT      651
                        Met Thr Ser His Tyr Val Ile Ala Ile Phe
                          1               5                  10

GCC CTG ATG AGC TTC TGT TTA GCC ACT GCA GGT CCA GAG CCT GGT GCA      699
Ala Leu Met Ser Phe Cys Leu Ala Thr Ala Gly Pro Glu Pro Gly Ala
                 15                  20                  25

CTG TGT GAA CTG TCA CCT GTC AGT GCC TCC CAT CCT GTC CAG GCC TTG      747
Leu Cys Glu Leu Ser Pro Val Ser Ala Ser His Pro Val Gln Ala Leu
             30                  35                  40

ATG GAG AGC TTC ACT GTT TTG TCA GGC TGT GCC AGC AGA GGC ACA ACT      795
Met Glu Ser Phe Thr Val Leu Ser Gly Cys Ala Ser Arg Gly Thr Thr
         45                  50                  55

GGG CTG CCA CAG GAG GTG CAT GTC CTG AAT CTC GCA CTG CGC CAG GGG      843
Gly Leu Pro Gln Glu Val His Val Leu Asn Leu Ala Leu Arg Gln Gly
     60                  65                  70

CCT GGC CAG CTA CAG AGA GAG GTC ACA CTT CAC CTG AAT CCC ATC TCC      891
Pro Gly Gln Leu Gln Arg Glu Val Thr Leu His Leu Asn Pro Ile Ser
 75                  80                  85                  90

TCA GTC CAC ATC CAC CAC AAG TCT GTT GTG TTC CTG CTC AAC TCC CCA      939
Ser Val His Ile His His Lys Ser Val Val Phe Leu Leu Asn Ser Pro
                 95                 100                 105

CAC CCC CTG GTG TGG CAT CTG AAG ACA GAG AGA CTT GCC ACT GGG GTC      987
His Pro Leu Val Trp His Leu Lys Thr Glu Arg Leu Ala Thr Gly Val
            110                 115                 120
```

```
TCC AGA CTG TTT TTG GTG TCT GAG GGT TCT GTG GTC CAG TTT TCA TCA     1035
Ser Arg Leu Phe Leu Val Ser Glu Gly Ser Val Val Gln Phe Ser Ser
        125                 130                 135

GCA AAC TTC TCC TTG ACA GCA GAA ACA GAA GAA AGG AAC TTC CCC CAT     1083
Ala Asn Phe Ser Leu Thr Ala Glu Thr Glu Glu Arg Asn Phe Pro His
    140                 145                 150

GGA AAT GAA CAT CTG TTA AAT TGG GCC CGA AAA GAG TAT GGA GCA GTT     1131
Gly Asn Glu His Leu Leu Asn Trp Ala Arg Lys Glu Tyr Gly Ala Val
155                 160                 165                 170

ACT TCA TTC ACC GAA CTC AAG ATA GCA AGA AAC ATT TAT ATT AAA GTG     1179
Thr Ser Phe Thr Glu Leu Lys Ile Ala Arg Asn Ile Tyr Ile Lys Val
                175                 180                 185

GGG GAA GAT CAA GTG TTC CCT CCA AAG TGC AAC ATA GGG AAG AAT TTT     1227
Gly Glu Asp Gln Val Phe Pro Pro Lys Cys Asn Ile Gly Lys Asn Phe
            190                 195                 200

CTC TCA CTC AAT TAC CTT GCT GAG TAC CTT CAA CCC AAA GCA GCA GAA     1275
Leu Ser Leu Asn Tyr Leu Ala Glu Tyr Leu Gln Pro Lys Ala Ala Glu
        205                 210                 215

GGG TGT GTG ATG TCC AGC CAG CCC CAG AAT GAG GAA GTA CAC ATC ATC     1323
Gly Cys Val Met Ser Ser Gln Pro Gln Asn Glu Glu Val His Ile Ile
    220                 225                 230

GAG CTA ATC ACC CCC AAC TCT AAC CCC TAC AGT GCT TTC CAG GTG GAT     1371
Glu Leu Ile Thr Pro Asn Ser Asn Pro Tyr Ser Ala Phe Gln Val Asp
235                 240                 245                 250

ATA ACA ATT GAT ATA AGA CCT TCT CAA GAG GAT CTT GAA GTG GTC AAA     1419
Ile Thr Ile Asp Ile Arg Pro Ser Gln Glu Asp Leu Glu Val Val Lys
                255                 260                 265

AAT CTC ATC CTG ATC TTG AAG TGC AAA AAG TCT GTC AAC TGG GTG ATC     1467
Asn Leu Ile Leu Ile Leu Lys Cys Lys Lys Ser Val Asn Trp Val Ile
            270                 275                 280

AAA TCT TTT GAT GTT AAG GGA AGC CTG AAA ATT ATT GCT CCT AAC AGT     1515
Lys Ser Phe Asp Val Lys Gly Ser Leu Lys Ile Ile Ala Pro Asn Ser
        285                 290                 295

ATT GGC TTT GGA AAA GAG AGT GAA AGA TCT ATG ACA ATG ACC AAA TCA     1563
Ile Gly Phe Gly Lys Glu Ser Glu Arg Ser Met Thr Met Thr Lys Ser
    300                 305                 310

ATA AGA GAT GAC ATT CCT TCA ACC CAA GGG AAT CTG GTG AAG TGG GCT     1611
Ile Arg Asp Asp Ile Pro Ser Thr Gln Gly Asn Leu Val Lys Trp Ala
315                 320                 325                 330

TTG GAC AAT GGC TAT AGT CCA ATA ACT TCA TAC ACA ATG GCT CCT GTG     1659
Leu Asp Asn Gly Tyr Ser Pro Ile Thr Ser Tyr Thr Met Ala Pro Val
                335                 340                 345

GCA ATA GTA TTT CAT CTT CGG CTT GAA AAT AAT GAG GAG ATG GGA GAT     1707
Ala Ile Val Phe His Leu Arg Leu Glu Asn Asn Glu Glu Met Gly Asp
            350                 355                 360

GAG GAA GTC CAC ACT ATT CCT CCT GAG CTA CGG ATC CTG CTG GAC CCT     1755
Glu Glu Val His Thr Ile Pro Pro Glu Leu Arg Ile Leu Leu Asp Pro
        365                 370                 375

GGT GCC CTG CCT GCC CTG CAG AAC CCG CCC ATC CGG GGA GGG GAA GGC     1803
Gly Ala Leu Pro Ala Leu Gln Asn Pro Pro Ile Arg Gly Gly Glu Gly
    380                 385                 390
```

```
CAA AAT GGA GGC CTT CCG TTT CCT TTC CCA GAT ATT TCC AGG AGA GTC      1851
Gln Asn Gly Gly Leu Pro Phe Pro Phe Pro Asp Ile Ser Arg Arg Val
395                 400                 405                 410

TGG AAT GAA GAG GGA GAA GAT GGG CTC CCT CGG CCA AAG GAC CCT GTC      1899
Trp Asn Glu Glu Gly Glu Asp Gly Leu Pro Arg Pro Lys Asp Pro Val
                415                 420                 425

ATT CCC AGC ATA CAA CTG TTT CCT GGT CTC AGA GAG CCA GAA GAG GTG      1947
Ile Pro Ser Ile Gln Leu Phe Pro Gly Leu Arg Glu Pro Glu Glu Val
            430                 435                 440

CAA GGG AGC GTG GAT ATT GCC CTG TCT GTC AAA TGT GAC AAT GAG AAG      1995
Gln Gly Ser Val Asp Ile Ala Leu Ser Val Lys Cys Asp Asn Glu Lys
        445                 450                 455

ATG ATC GTG GCT GTA GAA AAA GAT TCT TTT CAG GCC AGT GGC TAC TCG      2043
Met Ile Val Ala Val Glu Lys Asp Ser Phe Gln Ala Ser Gly Tyr Ser
    460                 465                 470

GGG ATG GAC GTC ACC CTG TTG GAT CCT ACC TGC AAG GCC AAG ATG AAT      2091
Gly Met Asp Val Thr Leu Leu Asp Pro Thr Cys Lys Ala Lys Met Asn
475                 480                 485                 490

GGC ACA CAC TTT GTT TTG GAG TCT CCT CTG AAT GGC TGC GGT ACT CGG      2139
Gly Thr His Phe Val Leu Glu Ser Pro Leu Asn Gly Cys Gly Thr Arg
                495                 500                 505

CCC CGG TGG TCA GCC CTT GAT GGT GTG GTC TAC TAT AAC TCC ATT GTG      2187
Pro Arg Trp Ser Ala Leu Asp Gly Val Val Tyr Tyr Asn Ser Ile Val
            510                 515                 520

ATA CAG GTT CCA GCC CTT GGG GAC AGT AGT GGT TGG CCA GAT GGT TAT      2235
Ile Gln Val Pro Ala Leu Gly Asp Ser Ser Gly Trp Pro Asp Gly Tyr
        525                 530                 535

GAA GAT CTG GAG TCA GGT GAT AAT GGA TTT CCG GGA GAT ATG GAT GAA      2283
Glu Asp Leu Glu Ser Gly Asp Asn Gly Phe Pro Gly Asp Met Asp Glu
    540                 545                 550

GGA GAT GCT TCC CTG TTC ACC CGA CCT GAA ATC GTG GTG TTT AAT TGC      2331
Gly Asp Ala Ser Leu Phe Thr Arg Pro Glu Ile Val Val Phe Asn Cys
555                 560                 565                 570

AGC CTT CAG CAG GTG AGG AAC CCC AGC AGC TTC CAG GAA CAG CCC CAC      2379
Ser Leu Gln Gln Val Arg Asn Pro Ser Ser Phe Gln Glu Gln Pro His
                575                 580                 585

GGA AAC ATC ACC TTC AAC ATG GAG CTA TAC AAC ACT GAC CTC TTT TTG      2427
Gly Asn Ile Thr Phe Asn Met Glu Leu Tyr Asn Thr Asp Leu Phe Leu
            590                 595                 600

GTG CCC TCC CAG GGC GTC TTC TCT GTG CCA GAG AAT GGA CAC GTT TAT      2475
Val Pro Ser Gln Gly Val Phe Ser Val Pro Glu Asn Gly His Val Tyr
        605                 610                 615

GTT GAG GTA TCT GTT ACT AAG GCT GAA CAA GAA CTG GGA TTT GCC ATC      2523
Val Glu Val Ser Val Thr Lys Ala Glu Gln Glu Leu Gly Phe Ala Ile
    620                 625                 630

CAA ACG TGC TTT ATC TCT CCA TAT TCG AAC CCT GAT AGG ATG TCT CAT      2571
Gln Thr Cys Phe Ile Ser Pro Tyr Ser Asn Pro Asp Arg Met Ser His
635                 640                 645                 650

TAC ACC ATT ATT GAG AAT ATT TGT CCT AAA GAT GAA TCT GTG AAA TTC      2619
Tyr Thr Ile Ile Glu Asn Ile Cys Pro Lys Asp Glu Ser Val Lys Phe
                655                 660                 665
```

```
TAC AGT CCC AAG AGA GTG CAC TTC CCT ATC CCG CAA GCT GAC ATG GAT        2667
Tyr Ser Pro Lys Arg Val His Phe Pro Ile Pro Gln Ala Asp Met Asp
            670                 675                 680

AAG AAG CGA TTC AGC TTT GTC TTC AAG CCT GTC TTC AAC ACC TCA CTG        2715
Lys Lys Arg Phe Ser Phe Val Phe Lys Pro Val Phe Asn Thr Ser Leu
        685                 690                 695

CTC TTT CTA CAG TGT GAG CTG ACG CTG TGT ACG AAG ATG GAG AAG CAC        2763
Leu Phe Leu Gln Cys Glu Leu Thr Leu Cys Thr Lys Met Glu Lys His
    700                 705                 710

CCC CAG AAG TTG CCT AAG TGT GTG CCT CCT GAC GAA GCC TGC ACC TCG        2811
Pro Gln Lys Leu Pro Lys Cys Val Pro Pro Asp Glu Ala Cys Thr Ser
715                 720                 725                 730

CTG GAC GCC TCG ATA ATC TGG GCC ATG ATG CAG AAT AAG AAG ACG TTC        2859
Leu Asp Ala Ser Ile Ile Trp Ala Met Met Gln Asn Lys Lys Thr Phe
                735                 740                 745

ACC AAG CCC CTT GCT GTG ATC CAC CAT GAA GCA GAA TCT AAA GAA AAA        2907
Thr Lys Pro Leu Ala Val Ile His His Glu Ala Glu Ser Lys Glu Lys
            750                 755                 760

GGT CCA AGC ATG AAG GAA CCA AAT CCA ATT TCT CCA CCA ATT TTC CAT        2955
Gly Pro Ser Met Lys Glu Pro Asn Pro Ile Ser Pro Pro Ile Phe His
        765                 770                 775

GGT CTG GAC ACC CTA ACC GTG ATG GGC ATT GCG TTT GCA GCC TTT GTG        3003
Gly Leu Asp Thr Leu Thr Val Met Gly Ile Ala Phe Ala Ala Phe Val
    780                 785                 790

ATC GGA GCA CTC CTG ACG GGG GCC TTG TGG TAC ATC TAT TCT CAC ACA        3051
Ile Gly Ala Leu Leu Thr Gly Ala Leu Trp Tyr Ile Tyr Ser His Thr
795                 800                 805                 810

GGG GAG ACA GCA GGA AGG CAG CAA GTC CCC ACC TCC CCG CCA GCC TCG        3099
Gly Glu Thr Ala Gly Arg Gln Gln Val Pro Thr Ser Pro Pro Ala Ser
                815                 820                 825

GAA AAC AGC AGT GCT GCC CAC AGC ATC GGC AGC ACG CAG AGC ACG CCT        3147
Glu Asn Ser Ser Ala Ala His Ser Ile Gly Ser Thr Gln Ser Thr Pro
            830                 835                 840

TGC TCC AGC AGC AGC ACG GCC TAGCCCAACC CAGGCCCAAC CCGGCCCAAC           3198
Cys Ser Ser Ser Ser Thr Ala
        845

CCAGCCCAGC CCAGCTCAGC TCAGCTACTC CAAGGGCAGG ACCAATGGCT GAGCCTCGTG    3258

TCCAGACTCA GAGGGCTGGA TTTTGGTTCC CTTGTAAAGA CAGAGTGAAT TTCAGTATAA    3318

AGATCACCCG TTGTATTCAC CCCACACCCA GGGCTAGTAT AAACATGACC CTGGGCTTCT    3378

GTACCACACT AGAATTCATG TGAGAAAGCT AAAATGGTGG TCTTCTCCAC CAGCCCCTCA    3438

CAGGCTTGGG GGTTTTCTAT GTGAAACACA TGCCAGTTTT TAAAATGCTG CTTTGTCCAG    3498

GTGAGAACAT CCATAATTTG GGGCCCTGAG TTTTACCCAG ACTCAAGGAG TTGGTAAAGG    3558

GTTAATAGCC AGATAGTAGA ACCAGTGAGG AGATGCGGCC AAAGATTCTT TATATCTGAA    3618

CCAAGATGTA AAACAAGAAA TGCTTTGAGG CTTTCTAAGC GATCCTCCTG TCTAATTTGC    3678

ACCTTTGTCT GGATGCACTC TTCTGACCTT GCTGCCACAA CCTGTGGGGT CTGATGTGTC    3738

CCAAGATGGG TGCTGCCCTC AGGGACTGCA CCCTGACAAG TGTTAAGGCA ACATTCCTTG    3798
```

CTTGTGCCCT GGGCCAAAAC CAATGCTGAT GACCTTATCA GCTTCCTGTT TCTTCCCATA 3858

CTGCATACAC CACTGCAAAA TGTCTTAATG CAAATTTTGT ATTTCTTACA GGCCTACAGA 3918

AATTGAAAAT GACCAAAATC AGGAACCACA GATTTGTGCC CATTCCTAAT ATTTTGTTCT 3978

GCAAATTAAT GTATAATTTG AGGTGAAATT CAGTTATAAA GTCAAGGACG AATTTGCACA 4038

GTGATATATT TCTATGTGTA TGCAAGTACA AGTATATAAT ATGTCACCTG GCACATTCAT 4098

TTTCTCAGTT GAAGAAGAGA AAATTTGAAA ATGTCCTTAT GCTTTTAGAG TTGCAACTTA 4158

AGTATATTTG GTAGGGTGAG TGTTTCCACT CAAAATATGT CAACTTAAAA AAAAA 4213

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 849 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Thr Ser His Tyr Val Ile Ala Ile Phe Ala Leu Met Ser Phe Cys
1 5 10 15

Leu Ala Thr Ala Gly Pro Glu Pro Gly Ala Leu Cys Glu Leu Ser Pro
20 25 30

Val Ser Ala Ser His Pro Val Gln Ala Leu Met Glu Ser Phe Thr Val
35 40 45

Leu Ser Gly Cys Ala Ser Arg Gly Thr Thr Gly Leu Pro Gln Glu Val
50 55 60

His Val Leu Asn Leu Ala Leu Arg Gln Gly Pro Gly Gln Leu Gln Arg
65 70 75 80

Glu Val Thr Leu His Leu Asn Pro Ile Ser Ser Val His Ile His His
85 90 95

Lys Ser Val Val Phe Leu Leu Asn Ser Pro His Pro Leu Val Trp His
100 105 110

Leu Lys Thr Glu Arg Leu Ala Thr Gly Val Ser Arg Leu Phe Leu Val
115 120 125

Ser Glu Gly Ser Val Val Gln Phe Ser Ser Ala Asn Phe Ser Leu Thr
130 135 140

Ala Glu Thr Glu Glu Arg Asn Phe Pro His Gly Asn Glu His Leu Leu
145 150 155 160

Asn Trp Ala Arg Lys Glu Tyr Gly Ala Val Thr Ser Phe Thr Glu Leu
165 170 175

Lys Ile Ala Arg Asn Ile Tyr Ile Lys Val Gly Glu Asp Gln Val Phe
180 185 190

Pro Pro Lys Cys Asn Ile Gly Lys Asn Phe Leu Ser Leu Asn Tyr Leu
195 200 205

Ala Glu Tyr Leu Gln Pro Lys Ala Ala Glu Gly Cys Val Met Ser Ser
210 215 220

Gln Pro Gln Asn Glu Glu Val His Ile Ile Glu Leu Ile Thr Pro Asn
225 230 235 240

Ser Asn Pro Tyr Ser Ala Phe Gln Val Asp Ile Thr Ile Asp Ile Arg
245 250 255

Pro Ser Gln Glu Asp Leu Glu Val Val Lys Asn Leu Ile Leu Ile Leu
260 265 270

Lys Cys Lys Lys Ser Val Asn Trp Val Ile Lys Ser Phe Asp Val Lys
275 280 285

Gly Ser Leu Lys Ile Ile Ala Pro Asn Ser Ile Gly Phe Gly Lys Glu
290 295 300

Ser Glu Arg Ser Met Thr Met Thr Lys Ser Ile Arg Asp Asp Ile Pro
305 310 315 320

Ser Thr Gln Gly Asn Leu Val Lys Trp Ala Leu Asp Asn Gly Tyr Ser
325 330 335

Pro Ile Thr Ser Tyr Thr Met Ala Pro Val Ala Ile Val Phe His Leu
340 345 350

Arg Leu Glu Asn Asn Glu Glu Met Gly Asp Glu Glu Val His Thr Ile
355 360 365

Pro Pro Glu Leu Arg Ile Leu Leu Asp Pro Gly Ala Leu Pro Ala Leu
370 375 380

Gln Asn Pro Pro Ile Arg Gly Gly Glu Gly Gln Asn Gly Gly Leu Pro
385 390 395 400

Phe Pro Phe Pro Asp Ile Ser Arg Arg Val Trp Asn Glu Glu Gly Glu
405 410 415

Asp Gly Leu Pro Arg Pro Lys Asp Pro Val Ile Pro Ser Ile Gln Leu
420 425 430

Phe Pro Gly Leu Arg Glu Pro Glu Glu Val Gln Gly Ser Val Asp Ile
435 440 445

Ala Leu Ser Val Lys Cys Asp Asn Glu Lys Met Ile Val Ala Val Glu
450 455 460

Lys Asp Ser Phe Gln Ala Ser Gly Tyr Ser Gly Met Asp Val Thr Leu
465 470 475 480

Leu Asp Pro Thr Cys Lys Ala Lys Met Asn Gly Thr His Phe Val Leu
485 490 495

Glu Ser Pro Leu Asn Gly Cys Gly Thr Arg Pro Arg Trp Ser Ala Leu
500 505 510

Asp Gly Val Val Tyr Tyr Asn Ser Ile Val Ile Gln Val Pro Ala Leu
515 520 525

Gly Asp Ser Ser Gly Trp Pro Asp Gly Tyr Glu Asp Leu Glu Ser Gly
530 535 540

Asp Asn Gly Phe Pro Gly Asp Met Asp Glu Gly Asp Ala Ser Leu Phe
545 550 555 560

Thr Arg Pro Glu Ile Val Val Phe Asn Cys Ser Leu Gln Gln Val Arg
565 570 575

```
Asn Pro Ser Ser Phe Gln Glu Gln Pro His Gly Asn Ile Thr Phe Asn
            580                 585                 590

Met Glu Leu Tyr Asn Thr Asp Leu Phe Leu Val Pro Ser Gln Gly Val
        595                 600                 605

Phe Ser Val Pro Glu Asn Gly His Val Tyr Val Glu Val Ser Val Thr
    610                 615                 620

Lys Ala Glu Gln Glu Leu Gly Phe Ala Ile Gln Thr Cys Phe Ile Ser
625                 630                 635                 640

Pro Tyr Ser Asn Pro Asp Arg Met Ser His Tyr Thr Ile Ile Glu Asn
                645                 650                 655

Ile Cys Pro Lys Asp Glu Ser Val Lys Phe Tyr Ser Pro Lys Arg Val
            660                 665                 670

His Phe Pro Ile Pro Gln Ala Asp Met Asp Lys Lys Arg Phe Ser Phe
        675                 680                 685

Val Phe Lys Pro Val Phe Asn Thr Ser Leu Leu Phe Leu Gln Cys Glu
    690                 695                 700

Leu Thr Leu Cys Thr Lys Met Glu Lys His Pro Gln Lys Leu Pro Lys
705                 710                 715                 720

Cys Val Pro Pro Asp Glu Ala Cys Thr Ser Leu Asp Ala Ser Ile Ile
                725                 730                 735

Trp Ala Met Met Gln Asn Lys Lys Thr Phe Thr Lys Pro Leu Ala Val
            740                 745                 750

Ile His His Glu Ala Glu Ser Lys Glu Lys Gly Pro Ser Met Lys Glu
        755                 760                 765

Pro Asn Pro Ile Ser Pro Pro Ile Phe His Gly Leu Asp Thr Leu Thr
    770                 775                 780

Val Met Gly Ile Ala Phe Ala Ala Phe Val Ile Gly Ala Leu Leu Thr
785                 790                 795                 800

Gly Ala Leu Trp Tyr Ile Tyr Ser His Thr Gly Glu Thr Ala Gly Arg
                805                 810                 815

Gln Gln Val Pro Thr Ser Pro Pro Ala Ser Glu Asn Ser Ser Ala Ala
            820                 825                 830

His Ser Ile Gly Ser Thr Gln Ser Thr Pro Cys Ser Ser Ser Ser Thr
        835                 840                 845

Ala
```

We claim:

1. A polypeptide of at least 155 amino acids that binds to TGF-β and that has a sequence consisting essentially of a sequence from a portion of a mammalian betaglycan within about one-third of the extracellular domain closest to the cell membrane.

2. The polypeptide of claim 1 wherein the portion of a mammalian betaglycan is about one-fifth of the extracellular domain of a mammalian betaglycan closest to the cell membrane.

3. The polypeptide of claim 1 wherein the portion of a mammalian betaglycan is about one-fourth of the extracellular domain of a mammalian betaglycan closest to the cell membrane.

4. The polypeptide of claim 1 wherein the mammalian betaglycan is from a human, rat or pig.

5. The polypeptide of claim 1 wherein the sequence consists essentially of at least amino acids 615 to 769 of SEQ ID NO:2 to at most amino acids 501 to 769 of SEQ ID NO:2.

6. The polypeptide of claim 1 wherein the sequence consists essentially of at least amino acids 543 to 769 of SEQ ID NO:2 to at most amino acids 501 to 769 of SEQ ID NO:2.

7. The polypeptide of claim 1 wherein the sequence consists essentially of amino acids 543 to 769 of SEQ ID NO:2.

8. A soluble polypeptide having the formula A-B-C, wherein A is a sequence excluding amino acid sequences of more than 4 amino acids from amino acids 1 to 543 of SEQ ID NO: 2; B is an amino acid sequence consisting essentially of at least amino acids 543 to 769 of SEQ ID NO:2 to at most amino acids 501 to 769 of SEQ ID NO:2; and C is an amino acid sequence.

9. The polypeptide of claim 8 wherein A and C have no more than 100 amino acids.

10. A soluble polypeptide having the formula A-B-C, wherein A is an amino acid sequence excluding sequences of more than 4 amino acids from amino acids 1 to 543 of SEQ ID NO 2; B is at least 155 amino acids in a sequence consisting essentially of an amino acid sequence within amino acids 543 to 769 of SEQ ID NO:2; and C is an amino acid sequence.

11. The polypeptide of claim 10 wherein A and C have no more than 100 amino acids.

12. An isolated nucleic acid molecule encoding any of the polypeptides of claims 1 to 11.

13. An expression vector comprising an expression control sequence operatively linked to a nucleic acid of claim 12.

14. A prokaryotic or eukaryotic cell transfected with an expression vector of claim 13 and capable of expressing a nucleic acid of claim 12.

15. A method of detecting TGF-β in a sample comprising contacting the sample with a polypeptide of claim 1 and determining the amount of TGF-β bound to the polypeptide.

16. A method of isolating TGF-β from a sample comprising contacting the sample with a polypeptide of claim 1 bound to a solid support to allow binding of TGF-β to the polypeptide and isolating the TGF-β from the polypeptide.

17. A method of enhancing the binding of TGF-β to type II receptor comprising contacting a cell bearing a type II receptor with TGF-β and a polypeptide of claim 1.

18. A method of enhancing suppression of cell growth by TGF-β comprising contacting a cell with TGF-β and a polypeptide of claim 1.

19. A pharmaceutical composition comprising a polypeptide according to claim 1, 8 or 10 in a pharmaceutically acceptable carrier.

20. The pharmaceutical composition of claim 19 further comprising TGF-β.

21. A method of treating a subject with a condition ameliorated by the enhanced binding of TGF-β to type II receptor or by the enhanced suppression of cell growth by TGF-β comprising administering a therapeutically effective amount of a pharmaceutical composition having a polypeptide of claim 1.

22. The method of claim 21 wherein the pharmaceutical composition further comprises TGF-β.

23. A decoy betaglycan polypeptide.

24. The decoy betaglycan polypeptide of claim 23 having a disabling modification within about one-fourth of the extracellular domain of a mammalian betaglycan closest to the cell membrane.

25. The decoy betaglycan polypeptide of claim 24 having a disabling modification within the sequence of amino acids 543 to 769 of SEQ ID NO:2.

26. A method of treating a subject with a condition ameliorated by the diminished binding of TGF-β to a TGF-β receptor or by the inhibition of the suppression of cell growth by TGF-β comprising administering to the subject a therapeutically effective amount of a pharmaceutical composition having a decoy betaglycan polypeptide.

27. A method of suppressing TGF-β-induced deposition of extracellular matrix in a subject comprising administering to the subject a therapeutically effective amount of a pharmaceutical composition having a decoy betaglycan polypeptide.

28. An anti-betaglycan-binding-site antibody.

29. A method of treating a subject with a condition ameliorated by the diminished binding of TGF-β to a TGF-β receptor or by the inhibition of the suppression of cell growth by TGF-β comprising administering to the subject a therapeutically effective amount of a pharmaceutical composition having an anti-betaglycan-binding-site antibody of claim 28.

30. A method of suppressing TGF-β-induced deposition of extracellular matrix in a subject comprising administering to the subject a therapeutically effective amount of a pharmaceutical composition having an anti-betaglycan-binding-site antibody of claim 28.

31. An anti-idiotypic antibody that mimics the ability of a polypeptide of claim 1 to bind TGF-β but that lacks the biological function of enhancing binding of TGF-β to a TGF-β receptor or suppressing cell growth.

32. A method of treating a subject with a condition ameliorated by the diminished binding of TGF-β to a TGF-β receptor or by the inhibition of the suppression of cell growth by TGF-β comprising administering to the subject a therapeutically effective amount of a pharmaceutical composition having an anti-idiotypic antibody of claim 31.

33. A method of suppressing TGF-β-induced deposition of extracellular matrix in a subject comprising administering to the subject a therapeutically effective amount of a pharmaceutical composition having an anti-idiotypic antibody of claim 31.

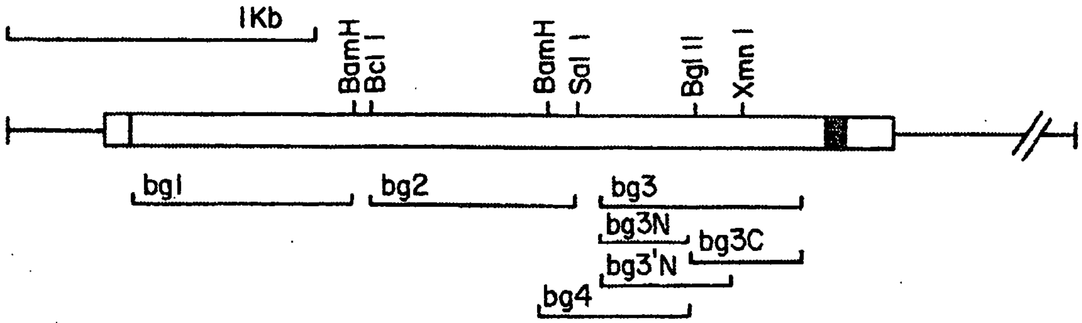


FIG. 1A

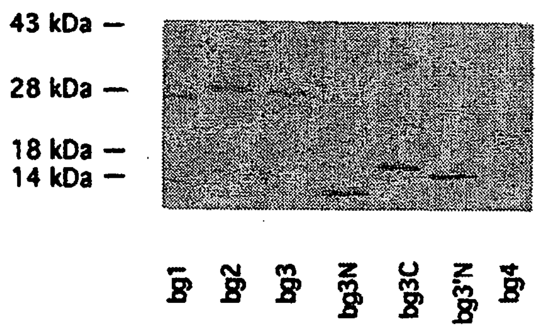


FIG. 1B

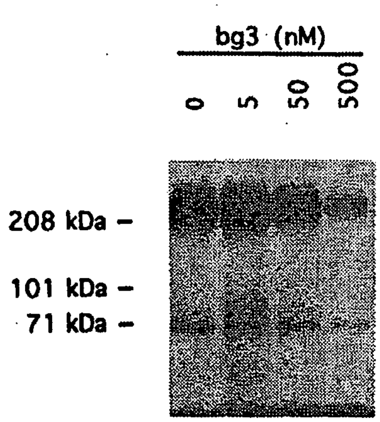


FIG. 3

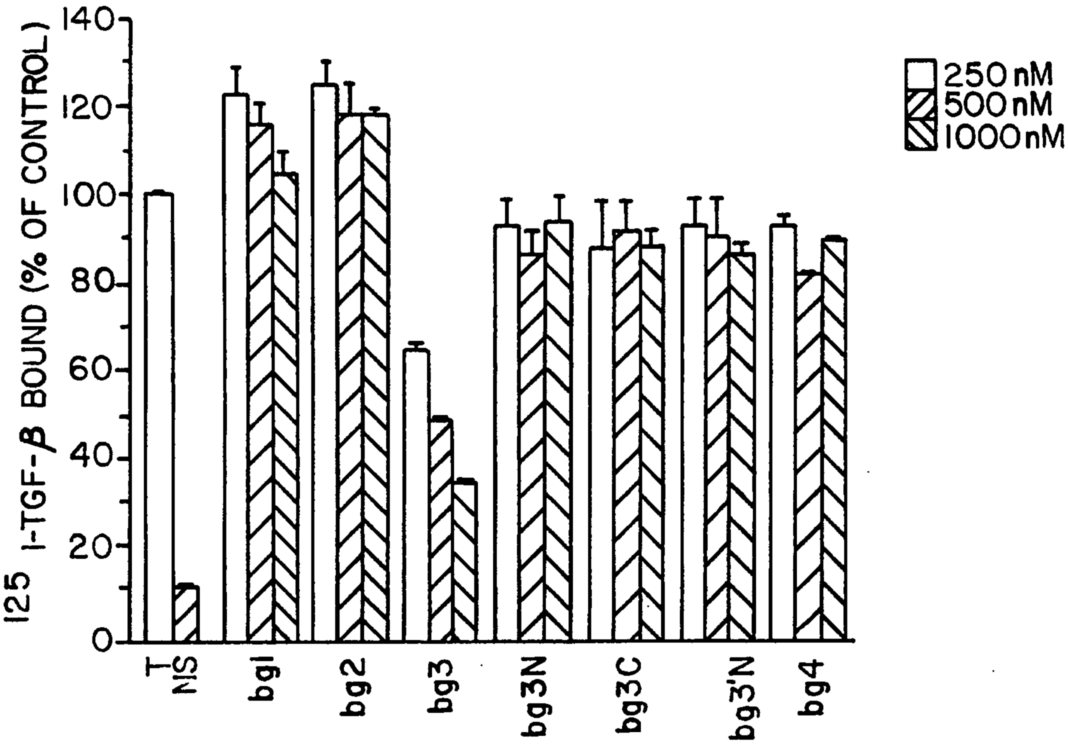

125 I-TGF-β BOUND (% OF CONTROL)
140
120
100
80
60
40
20
0
250 nM
500 nM
1000nM
T
NS
bg1
bg2
bg3
bg3N
bg3C
bg3'N
bg4

FIG. 2A

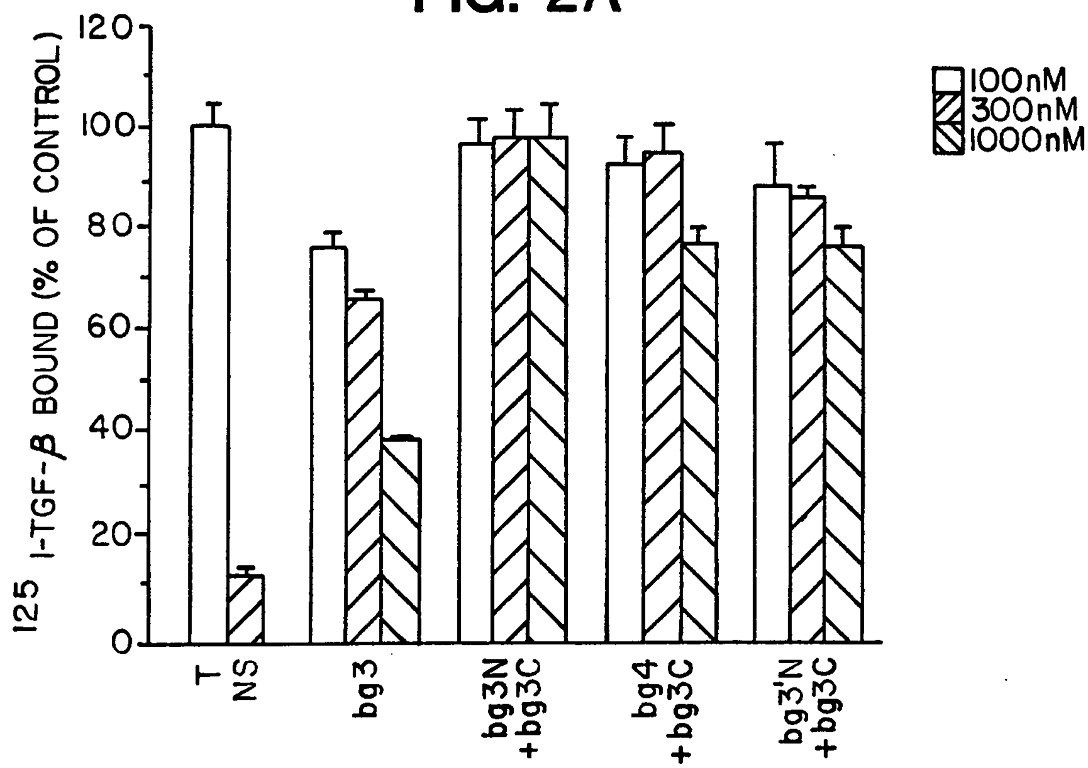

125 I-TGF-β BOUND (% OF CONTROL)
120
100
80
60
40
20
0
100nM
300nM
1000nM
T
NS
bg3
bg3N +bg3C
bg4 +bg3C
bg3'N +bg3C

FIG. 2B

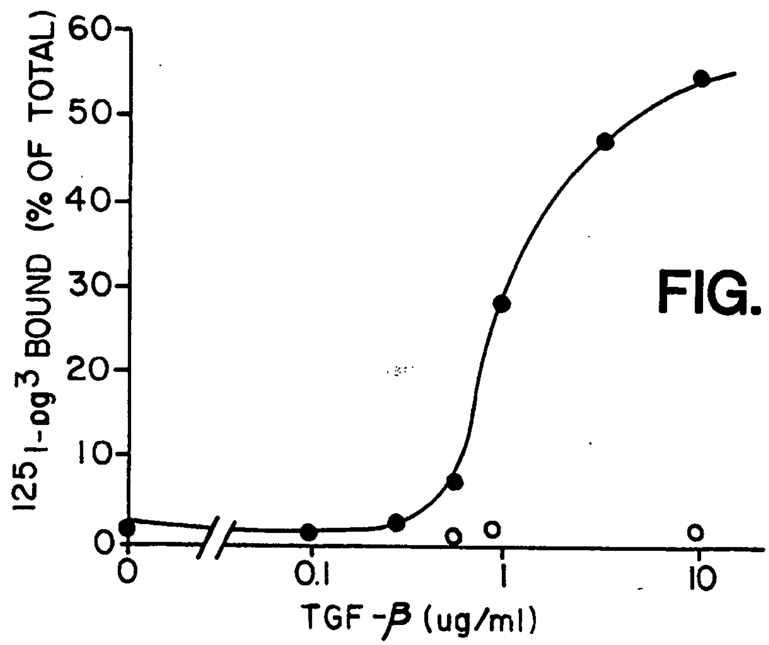

125I-bg3 BOUND (% OF TOTAL)
60
50
40
30
20
10
0
0
0.1
1
10
TGF-β (ug/ml)

FIG. 4A

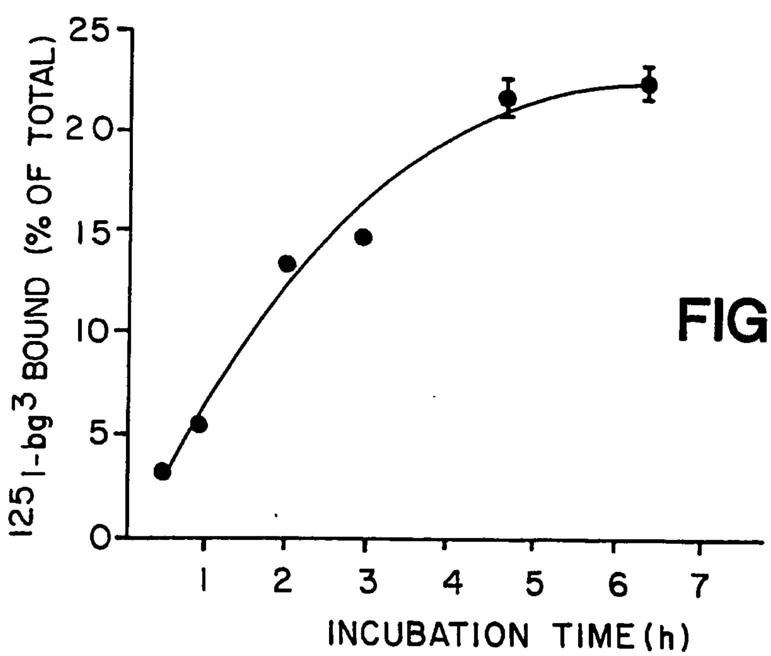

125I-bg3 BOUND (% OF TOTAL)
25
20
15
10
5
0
1
2
3
4
5
6
7
INCUBATION TIME (h)

FIG. 4B

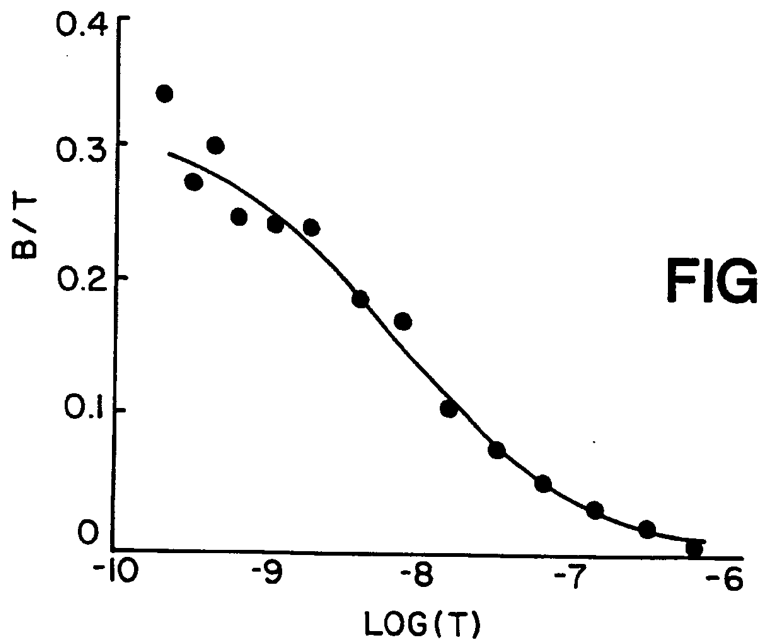


FIG. 5A

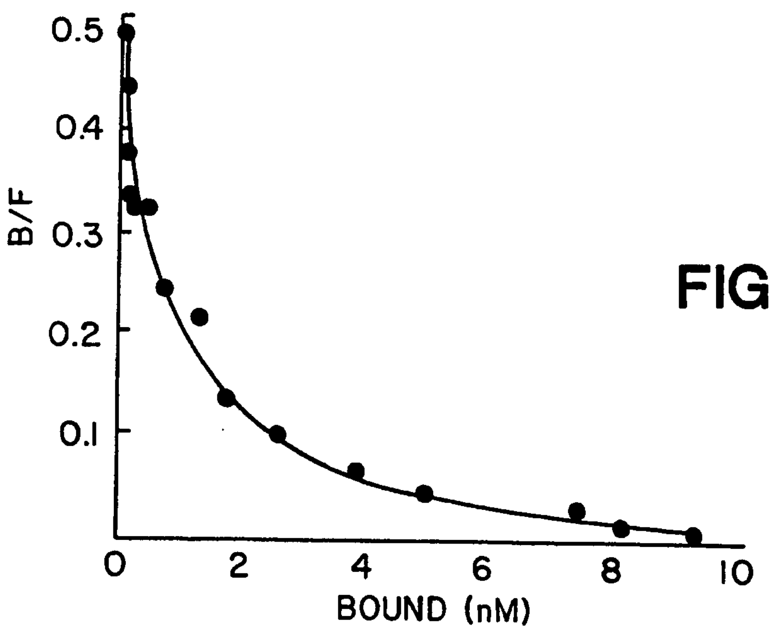


FIG. 5B

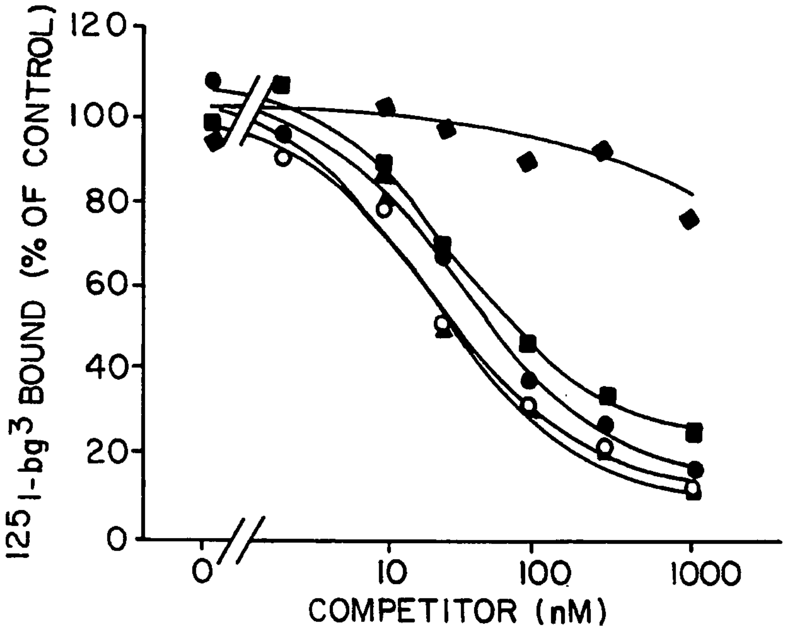

120
100
80
60
40
20
0
^{125}I-bg^3 BOUND (% OF CONTROL)
0
10
100
1000
COMPETITOR (nM)

FIG. 6

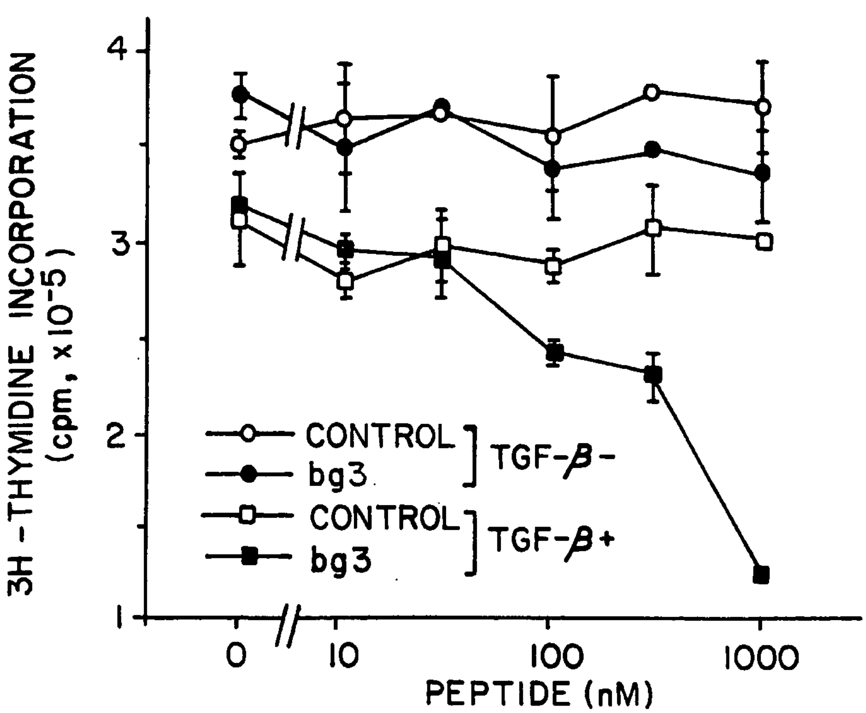
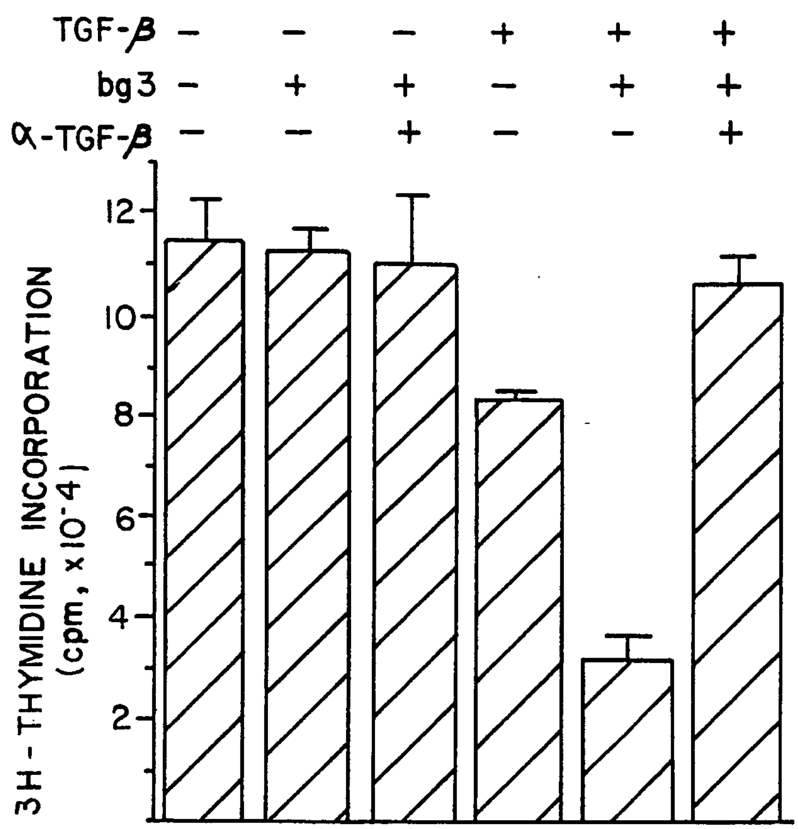

4
3
2
1
3H -THYMIDINE INCORPORATION (cpm, x10^{-5})
CONTROL
bg3
TGF-β-
CONTROL
bg3
TGF-β+
0
10
100
1000
PEPTIDE (nM)

FIG. 7



TGF-β − − − + + +
bg3 − + + − + +
α-TGF-β − − + − − +
3H - THYMIDINE INCORPORATION (cpm, x10^-4)
12
10
8
6
4
2

FIG. 8

-240
CAGGAGGTGA AAGTCCCCGG CGGTCCGGAT GGCGCAGTTG CACTGCGCTG CTGAGCTCGC
GGCCGCCTGC GCACACTGGG GGGACTCGCT TCGGCTAGTA ACTCCTCCAC CTCGCGGCGG
ACGACCGGTC CTGGACACGC TGCCTGCGAG GCAAGTTGAA CAGTGCAGAG AAGGATCTTA
AAGCTACACC CGACTTGCCA CGATTGCCTT CAATCTGAAG AACCAAAGGC TGTTGGAGAG -1

```
ATG GCA GTG ACA TCC CAC CAC ATG ATC CCG GTG ATG GTT GTC CTG ATG    48
Met Ala Val Thr Ser His His Met Ile Pro Val Met Val Val Leu Met
  1               5                  10                  15

AGC GCC TGC CTG GCC ACC GCC GGT CCA GAG CCC AGC ACC CGG TGT GAA    96
Ser Ala Cys Leu Ala Thr Ala Gly Pro Glu Pro Ser Thr Arg Cys Glu
             20                  25                  30

CTG TCA CCA ATC AAC GCC TCT CAC CCA GTC CAG GCC TTG ATG GAG AGC   144
Leu Ser Pro Ile Asn Ala Ser His Pro Val Gln Ala Leu Met Glu Ser
         35                  40                  45

TTC ACC GTT CTG TCT GGC TGT GCC AGC AGA GGC ACC ACC GGG CTG CCA   192
Phe Thr Val Leu Ser Gly Cys Ala Ser Arg Gly Thr Thr Gly Leu Pro
     50                  55                  60

AGG GAG GTC CAT GTC CTA AAC CTC CGA AGT ACA GAT CAG GGA CCA GGC   240
Arg Glu Val His Val Leu Asn Leu Arg Ser Thr Asp Gln Gly Pro Gly
 65                  70                  75                  80

CAG CGG CAG AGA GAG GTT ACC CTG CAC CTG AAC CCC ATT GCC TCG GTG   288
Gln Arg Gln Arg Glu Val Thr Leu His Leu Asn Pro Ile Ala Ser Val
                 85                  90                  95

CAC ACT CAC CAC AAA CCT ATC GTG TTC CTG CTC AAC TCC CCC CAG CCC   336
His Thr His His Lys Pro Ile Val Phe Leu Leu Asn Ser Pro Gln Pro
            100                 105                 110

CTG GTG TGG CAT CTG AAG ACG GAG AGA CTG GCC GCT GGT GTC CCC AGA   384
Leu Val Trp His Leu Lys Thr Glu Arg Leu Ala Ala Gly Val Pro Arg
        115                 120                 125

CTC TTC CTG GTT TCG GAG GGT TCT GTG GTC CAG TTT CCA TCA GGA AAC   432
Leu Phe Leu Val Ser Glu Gly Ser Val Val Gln Phe Pro Ser Gly Asn
    130                 135                 140

TTC TCC TTG ACA GCA GAA ACA GAG GAA AGG AAT TTC CCT CAA GAA AAT   480
Phe Ser Leu Thr Ala Glu Thr Glu Glu Arg Asn Phe Pro Gln Glu Asn
145                 150                 155                 160

GAA CAT CTC GTG CGC TGG GCC CAA AAG GAA TAT GGA GCA GTG ACT TCG   528
Glu His Leu Val Arg Trp Ala Gln Lys Glu Tyr Gly Ala Val Thr Ser
                165                 170                 175
```

FIG. 9A

```
TTC ACT GAA CTC AAG ATA GCA AGA AAC ATC TAT ATT AAA GTG GGA GAA    576
Phe Thr Glu Leu Lys Ile Ala Arg Asn Ile Tyr Ile Lys Val Gly Glu
            180                 185                 190

GAT CAA GTG TTT CCT CCT ACG TGT AAC ATA GGG AAG AAT TTC CTC TCA    624
Asp Gln Val Phe Pro Pro Thr Cys Asn Ile Gly Lys Asn Phe Leu Ser
        195                 200                 205

CTC AAT TAC CTT GCC GAG TAC CTT CAA CCC AAA GCC GCC GAA GGT TGT    672
Leu Asn Tyr Leu Ala Glu Tyr Leu Gln Pro Lys Ala Ala Glu Gly Cys
    210                 215                 220

GTC CTG CCC AGT CAG CCC CAT GAA AAG GAA GTA CAC ATC ATC GAG TTA    720
Val Leu Pro Ser Gln Pro His Glu Lys Glu Val His Ile Ile Glu Leu
225                 230                 235                 240

ATT ACC CCC AGC TCG AAC CCT TAC AGC GCT TTC CAG GTG GAT ATA ATA    768
Ile Thr Pro Ser Ser Asn Pro Tyr Ser Ala Phe Gln Val Asp Ile Ile
                245                 250                 255

GTT GAC ATA CGA CCT GCT CAA GAG GAT CCC GAG GTG GTC AAA AAC CTT    816
Val Asp Ile Arg Pro Ala Gln Glu Asp Pro Glu Val Val Lys Asn Leu
            260                 265                 270

GTC CTG ATC TTG AAG TGC AAA AAG TCT GTC AAC TGG GTG ATC AAG TCT    864
Val Leu Ile Leu Lys Cys Lys Lys Ser Val Asn Trp Val Ile Lys Ser
        275                 280                 285

TTT GAC GTC AAG GGA AAC TTG AAA GTC ATT GCT CCC AAC AGT ATC GGC    912
Phe Asp Val Lys Gly Asn Leu Lys Val Ile Ala Pro Asn Ser Ile Gly
    290                 295                 300

TTT GGA AAA GAG AGT GAA CGA TCC ATG ACA ATG ACC AAA TTG GTA AGA    960
Phe Gly Lys Glu Ser Glu Arg Ser Met Thr Met Thr Lys Leu Val Arg
305                 310                 315                 320

GAT GAC ATC CCT TCC ACC CAA GAG AAT CTG ATG AAG TGG GCA CTG GAC   1008
Asp Asp Ile Pro Ser Thr Gln Glu Asn Leu Met Lys Trp Ala Leu Asp
                325                 330                 335

AAT GGC TAC AGG CCA GTG ACG TCA TAC ACA ATG GCT CCC GTG GCT AAT   1056
Asn Gly Tyr Arg Pro Val Thr Ser Tyr Thr Met Ala Pro Val Ala Asn
            340                 345                 350

AGA TTT CAT CTT CGG CTT GAG AAC AAC GAG GAG ATG AGA GAT GAG GAA   1104
Arg Phe His Leu Arg Leu Glu Asn Asn Glu Glu Met Arg Asp Glu Glu
        355                 360                 365
```

FIG. 9B

```
GTC CAC ACC ATT CCT CCT GAG CTT CGT ATC CTG CTG GAC CCT GAC CAC   1152
Val His Thr Ile Pro Pro Glu Leu Arg Ile Leu Leu Asp Pro Asp His
    370                 375                 380

CCG CCC GCC CTG GAC AAC CCA CTC TTC CCA GGA GAG GGA AGC CCA AAT   1200
Pro Pro Ala Leu Asp Asn Pro Leu Phe Pro Gly Glu Gly Ser Pro Asn
385                 390                 395                 400

GGT GGT CTC CCC TTT CCA TTC CCG GAT ATC CCC AGG AGA GGC TGG AAG   1248
Gly Gly Leu Pro Phe Pro Phe Pro Asp Ile Pro Arg Arg Gly Trp Lys
                405                 410                 415

GAG GGC GAA GAT AGG ATC CCC CGG CCA AAG CAG CCC ATC GTT CCC AGT   1296
Glu Gly Glu Asp Arg Ile Pro Arg Pro Lys Gln Pro Ile Val Pro Ser
            420                 425                 430

GTT CAA CTG CTT CCT GAC CAC CGA GAA CCA GAA GAA GTG CAA GGG GGC   1344
Val Gln Leu Leu Pro Asp His Arg Glu Pro Glu Glu Val Gln Gly Gly
        435                 440                 445

GTG GAC ATC GCC CTG TCA GTC AAA TGT GAC CAT GAA AAG ATG GTC GTG   1392
Val Asp Ile Ala Leu Ser Val Lys Cys Asp His Glu Lys Met Val Val
    450                 455                 460

GCT GTA GAC AAA GAC TCT TTC CAG ACC AAT GGC TAC TCA GGG ATG GAG   1440
Ala Val Asp Lys Asp Ser Phe Gln Thr Asn Gly Tyr Ser Gly Met Glu
465                 470                 475                 480

CTC ACC CTG TTG GAT CCT TCG TGT AAA GCC AAA ATG AAT GGT ACT CAC   1488
Leu Thr Leu Leu Asp Pro Ser Cys Lys Ala Lys Met Asn Gly Thr His
                485                 490                 495

TTT GTT CTC GAG TCT CCC CTG AAT GGC TGT GGT ACT CGA CAT CGG AGG   1536
Phe Val Leu Glu Ser Pro Leu Asn Gly Cys Gly Thr Arg His Arg Arg
            500                 505                 510

TCG ACC CCG GAT GGT GTG GTT TAC TAT AAC TCT ATT GTG GTG CAG GCT   1584
Ser Thr Pro Asp Gly Val Val Tyr Tyr Asn Ser Ile Val Val Gln Ala
        515                 520                 525

CCG TCC CCT GGG GAT AGC AGT GGC TGG CCT GAT GGC TAT GAA GAC TTG   1632
Pro Ser Pro Gly Asp Ser Ser Gly Trp Pro Asp Gly Tyr Glu Asp Leu
    530                 535                 540

GAG TCA GGC GAT AAT GGA TTT CCT GGA GAC GGG GAT GAA GGA GAA ACT   1680
Glu Ser Gly Asp Asn Gly Phe Pro Gly Asp Gly Asp Glu Gly Glu Thr
545                 550                 555                 560
```

FIG. 9C

```
GCC CCC CTG AGC CGA GCT GGA GTG GTG GTG TTT AAC TGC AGC TTG CGG   1728
Ala Pro Leu Ser Arg Ala Gly Val Val Val Phe Asn Cys Ser Leu Arg
                565                 570                 575

CAG CTG AGG AAT CCC AGT GGC TTC CAG GGC CAG CTC GAT GGA AAT GCT   1776
Gln Leu Arg Asn Pro Ser Gly Phe Gln Gly Gln Leu Asp Gly Asn Ala
            580                 585                 590

ACC TTC AAC ATG GAG CTG TAT AAC ACA GAC CTC TTT CTG GTG CCC TCC   1824
Thr Phe Asn Met Glu Leu Tyr Asn Thr Asp Leu Phe Leu Val Pro Ser
        595                 600                 605

CCA GGG GTC TTC TCT GTG GCA GAG AAC GAG CAT GTT TAT GTT GAG GTG   1872
Pro Gly Val Phe Ser Val Ala Glu Asn Glu His Val Tyr Val Glu Val
    610                 615                 620

TCT GTC ACC AAG GCT GAC CAA GAT CTG GGA TTC GCC ATC CAA ACC TGC   1920
Ser Val Thr Lys Ala Asp Gln Asp Leu Gly Phe Ala Ile Gln Thr Cys
625                 630                 635                 640

TTT CTC TCT CCA TAC TCC AAC CCA GAC AGA ATG TCT GAT TAC ACC ATC   1968
Phe Leu Ser Pro Tyr Ser Asn Pro Asp Arg Met Ser Asp Tyr Thr Ile
                645                 650                 655

ATC GAG AAC ATC TGT CCG AAA GAC GAC TCT GTG AAG TTC TAC AGC TCC   2016
Ile Glu Asn Ile Cys Pro Lys Asp Asp Ser Val Lys Phe Tyr Ser Ser
            660                 665                 670

AAG AGA GTG CAC TTT CCC ATC CCG CAT GCT GAG GTG GAC AAG AAG CGC   2064
Lys Arg Val His Phe Pro Ile Pro His Ala Glu Val Asp Lys Lys Arg
        675                 680                 685

TTC AGC TTC CTG TTC AAG TCT GTG TTC AAC ACC TCC CTG CTC TTC CTG   2122
Phe Ser Phe Leu Phe Lys Ser Val Phe Asn Thr Ser Leu Leu Phe Leu
    690                 695                 700

CAC TGC GAG TTG ACT CTG TGC TCC AGG AAG AAG GGC TCC CTG AAG CTG   2160
His Cys Glu Leu Thr Leu Cys Ser Arg Lys Lys Gly Ser Leu Lys Leu
705                 710                 715                 720

CCG AGG TGT GTC ACT CCT GAC GAC GCC TGC ACT TCT CTC GAT GCC ACC   2208
Pro Arg Cys Val Thr Pro Asp Asp Ala Cys Thr Ser Leu Asp Ala Thr
                725                 730                 735

ATG ATC TGG ACC ATG ATG CAG AAT AAG AAG ACA TTC ACC AAG CCC CTG   2256
Met Ile Trp Thr Met Met Gln Asn Lys Lys Thr Phe Thr Lys Pro Leu
            740                 745                 750
```

FIG. 9D

```
GCT GTG GTC CTC CAG GTA GAC TAT AAA GAA AAT GTT CCC AGC ACT AAG  2304
Ala Val Val Leu Gln Val Asp Tyr Lys Glu Asn Val Pro Ser Thr Lys
        755                 760                 765

GAT TCC AGT CCA ATT CCT CCT CCT CCT CCA CAG ATT TTC CAT GGC CTG  2352
Asp Ser Ser Pro Ile Pro Pro Pro Pro Pro Gln Ile Phe His Gly Leu
    770                 775                 780

GAC ACG CTC ACC GTG ATG GGC ATT GCA TTT GCA GCA TTT GTG ATC GGA  2400
Asp Thr Leu Thr Val Met Gly Ile Ala Phe Ala Ala Phe Val Ile Gly
785                 790                 795                 800

GCG CTC CTG ACG GGG GCC TTG TGG TAC ATC TAC TCC CAC ACA GGG GAG  2448
Ala Leu Leu Thr Gly Ala Leu Trp Tyr Ile Tyr Ser His Thr Gly Glu
                805                 810                 815

ACA GCA CGA AGG CAG CAA GTC CCT ACC TCG CCG CCA GCC TCG GAG AAC  2496
Thr Ala Arg Arg Gln Gln Val Pro Thr Ser Pro Pro Ala Ser Glu Asn
            820                 825                 830

AGC AGC GCG GCC CAC AGC ATC GGC AGC ACT CAG AGT ACC CCC TGC TCT  2544
Ser Ser Ala Ala His Ser Ile Gly Ser Thr Gln Ser Thr Pro Cys Ser
        835                 840                 845

AGC AGC AGC ACA GCC TAGGTGGACA GACAGACGCC CGCCCACCGC AGCCAGGGCA  2599
Ser Ser Ser Thr Ala
    850

GGGCCCGATG CCAGTGCTGC GTGTCCACAG TCAGAAGTCT TGATCTGGGC TCCCTGTAAA  2659

GAAAGAGTGA ATTTCAGTAT ACAGACAGCC AGTTCTACCC ACCCCTTACC ACGGCCCACA  2719

TAAATGTGAC CCTGGGCATC TGTCACACGA AAGCTAAGCT GGTGGCCTTC CCCACCAGCC  2779

CCTCGCAGGA TGGGGGTTTC AATGTGAAAC ATCTGCCAGT TTTGTTTTGT TTTTTTAATG  2839

CTGCTTTGTC CAGGTGTCCA AACATCCATC ATTTGGGGTG GTCTGTTTTA CAGAGTAAAG  2899

GAGGCGGTGA AGGGACGTCA GCTAGTGTGT AGAGCCAAGG GGAGACAGCT AGGATTCTCG  2959

CCTAGCTGAA CCAAGGTGTA AAATAGAAGA CACGCTCC                         2997
```

FIG. 9E

```
GCGGACGCGA TCGCAACCTC CGCTTCCCGG GTTCAAGGGA TTCTCCTGCC TCAGCCTCCT 60
GAGTAGCTGG GACTACAGGC GCACCACTAC ACCCGGCCAA TTTTTGTATT TATTTATTTT120
ATTTTATTTT ATTTTTTAGT AGAGACGGGG TTTACCATGT TGGCCAGAAC AGTCTCGATC180
TCCTAACCTC GTGATCCAAC GCCTCGGCTC CCAAAGTGTT GAGTTACAGG CGTGAGCCAC240
CGCGCCCGTC CAGAATTATC TTTAAAATTT ATTTCTTTAA GATTTGTAGC TACTAAGAAA300
GAAAGGAGCT TTTTTTCCTT GGGCCTTCAA ACTGAAAGAA CCGCATGAGC CTGACGGCGC360
ATGGTCTTAA CATCAGGCTG TGCAGGAAGA AGCTATCTGC AGATGGATGC CAGCACACAC420
AAGGAAGCAG AGCTCTGGCA ACATTGAGTC AAAGCAAGGA CACAACATCA GAGGGACGGC480
AGAGAATCCT TGTGTGTAGT CTTTGGTGGC AGTTTGAAAA TTGCAAGGAG GGACTTTAAG540
ACTACTTCTG ATTTGCAAAG ATGGTCTGTG CTCCGAGCAG GCTAAAGTGA CTGGACGAGA600
CGCACTGTTG GAGAAATAAA A ATG ACT TCC CAT TAT GTG ATT GCC ATC TTT   651
                        Met Thr Ser His Tyr Val Ile Ala Ile Phe
                         1               5                  10

GCC CTG ATG AGC TTC TGT TTA GCC ACT GCA GGT CCA GAG CCT GGT GCA   699
Ala Leu Met Ser Phe Cys Leu Ala Thr Ala Gly Pro Glu Pro Gly Ala
                15                  20                  25

CTG TGT GAA CTG TCA CCT GTC AGT GCC TCC CAT CCT GTC CAG GCC TTG   747
Leu Cys Glu Leu Ser Pro Val Ser Ala Ser His Pro Val Gln Ala Leu
            30                  35                  40

ATG GAG AGC TTC ACT GTT TTG TCA GGC TGT GCC AGC AGA GGC ACA ACT   795
Met Glu Ser Phe Thr Val Leu Ser Gly Cys Ala Ser Arg Gly Thr Thr
        45                  50                  55

GGG CTG CCA CAG GAG GTG CAT GTC CTG AAT CTC GCA CTG CGC CAG GGG   843
Gly Leu Pro Gln Glu Val His Val Leu Asn Leu Ala Leu Arg Gln Gly
    60                  65                  70

CCT GGC CAG CTA CAG AGA GAG GTC ACA CTT CAC CTG AAT CCC ATC TCC   891
Pro Gly Gln Leu Gln Arg Glu Val Thr Leu His Leu Asn Pro Ile Ser
 75                  80                  85                  90

TCA GTC CAC ATC CAC CAC AAG TCT GTT GTG TTC CTG CTC AAC TCC CCA   939
Ser Val His Ile His His Lys Ser Val Val Phe Leu Leu Asn Ser Pro
                95                 100                 105

CAC CCC CTG GTG TGG CAT CTG AAG ACA GAG AGA CTT GCC ACT GGG GTC   987
His Pro Leu Val Trp His Leu Lys Thr Glu Arg Leu Ala Thr Gly Val
            110                 115                 120

TCC AGA CTG TTT TTG GTG TCT GAG GGT TCT GTG GTC CAG TTT TCA TCA  1035
Ser Arg Leu Phe Leu Val Ser Glu Gly Ser Val Val Gln Phe Ser Ser
        125                 130                 135
```

FIG. 10A

```
GCA AAC TTC TCC TTG ACA GCA GAA ACA GAA GAA AGG AAC TTC CCC CAT  1083
Ala Asn Phe Ser Leu Thr Ala Glu Thr Glu Glu Arg Asn Phe Pro His
    140                 145                 150

GGA AAT GAA CAT CTG TTA AAT TGG GCC CGA AAA GAG TAT GGA GCA GTT  1131
Gly Asn Glu His Leu Leu Asn Trp Ala Arg Lys Glu Tyr Gly Ala Val
155                 160                 165                 170

ACT TCA TTC ACC GAA CTC AAG ATA GCA AGA AAC ATT TAT ATT AAA GTG  1179
Thr Ser Phe Thr Glu Leu Lys Ile Ala Arg Asn Ile Tyr Ile Lys Val
                175                 180                 185

GGG GAA GAT CAA GTG TTC CCT CCA AAG TGC AAC ATA GGG AAG AAT TTT  1227
Gly Glu Asp Gln Val Phe Pro Pro Lys Cys Asn Ile Gly Lys Asn Phe
            190                 195                 200

CTC TCA CTC AAT TAC CTT GCT GAG TAC CTT CAA CCC AAA GCA GCA GAA  1275
Leu Ser Leu Asn Tyr Leu Ala Glu Tyr Leu Gln Pro Lys Ala Ala Glu
        205                 210                 215

GGG TGT GTG ATG TCC AGC CAG CCC CAG AAT GAG GAA GTA CAC ATC ATC  1323
Gly Cys Val Met Ser Ser Gln Pro Gln Asn Glu Glu Val His Ile Ile
    220                 225                 230

GAG CTA ATC ACC CCC AAC TCT AAC CCC TAC AGT GCT TTC CAG GTG GAT  1371
Glu Leu Ile Thr Pro Asn Ser Asn Pro Tyr Ser Ala Phe Gln Val Asp
235                 240                 245                 250

ATA ACA ATT GAT ATA AGA CCT TCT CAA GAG GAT CTT GAA GTG GTC AAA  1419
Ile Thr Ile Asp Ile Arg Pro Ser Gln Glu Asp Leu Glu Val Val Lys
                255                 260                 265

AAT CTC ATC CTG ATC TTG AAG TGC AAA AAG TCT GTC AAC TGG GTG ATC  1467
Asn Leu Ile Leu Ile Leu Lys Cys Lys Lys Ser Val Asn Trp Val Ile
            270                 275                 280

AAA TCT TTT GAT GTT AAG GGA AGC CTG AAA ATT ATT GCT CCT AAC AGT  1515
Lys Ser Phe Asp Val Lys Gly Ser Leu Lys Ile Ile Ala Pro Asn Ser
        285                 290                 295

ATT GGC TTT GGA AAA GAG AGT GAA AGA TCT ATG ACA ATG ACC AAA TCA  1563
Ile Gly Phe Gly Lys Glu Ser Glu Arg Ser Met Thr Met Thr Lys Ser
    300                 305                 310

ATA AGA GAT GAC ATT CCT TCA ACC CAA GGG AAT CTG GTG AAG TGG GCT  1611
Ile Arg Asp Asp Ile Pro Ser Thr Gln Gly Asn Leu Val Lys Trp Ala
315                 320                 325                 330

TTG GAC AAT GGC TAT AGT CCA ATA ACT TCA TAC ACA ATG GCT CCT GTG  1659
Leu Asp Asn Gly Tyr Ser Pro Ile Thr Ser Tyr Thr Met Ala Pro Val
                335                 340                 345

GCA ATA GTA TTT CAT CTT CGG CTT GAA AAT AAT GAG GAG ATG GGA GAT  1707
Ala Ile Val Phe His Leu Arg Leu Glu Asn Asn Glu Glu Met Gly Asp
            350                 355                 360
```

FIG. 10B

GAG GAA GTC CAC ACT ATT CCT CCT GAG CTA CGG ATC CTG CTG GAC CCT 1755
Glu Glu Val His Thr Ile Pro Pro Glu Leu Arg Ile Leu Leu Asp Pro
365 370 375

GGT GCC CTG CCT GCC CTG CAG AAC CCG CCC ATC CGG GGA GGG GAA GGC 1803
Gly Ala Leu Pro Ala Leu Gln Asn Pro Pro Ile Arg Gly Gly Glu Gly
380 385 390

CAA AAT GGA GGC CTT CCG TTT CCT TTC CCA GAT ATT TCC AGG AGA GTC 1851
Gln Asn Gly Gly Leu Pro Phe Pro Phe Pro Asp Ile Ser Arg Arg Val
395 400 405 410

TGG AAT GAA GAG GGA GAA GAT GGG CTC CCT CGG CCA AAG GAC CCT GTC 1899
Trp Asn Glu Glu Gly Glu Asp Gly Leu Pro Arg Pro Lys Asp Pro Val
415 420 425

ATT CCC AGC ATA CAA CTG TTT CCT GGT CTC AGA GAG CCA GAA GAG GTG 1947
Ile Pro Ser Ile Gln Leu Phe Pro Gly Leu Arg Glu Pro Glu Glu Val
430 435 440

CAA GGG AGC GTG GAT ATT GCC CTG TCT GTC AAA TGT GAC AAT GAG AAG 1995
Gln Gly Ser Val Asp Ile Ala Leu Ser Val Lys Cys Asp Asn Glu Lys
445 450 455

ATG ATC GTG GCT GTA GAA AAA GAT TCT TTT CAG GCC AGT GGC TAC TCG 2043
Met Ile Val Ala Val Glu Lys Asp Ser Phe Gln Ala Ser Gly Tyr Ser
460 465 470

GGG ATG GAC GTC ACC CTG TTG GAT CCT ACC TGC AAG GCC AAG ATG AAT 2091
Gly Met Asp Val Thr Leu Leu Asp Pro Thr Cys Lys Ala Lys Met Asn
475 480 485 490

GGC ACA CAC TTT GTT TTG GAG TCT CCT CTG AAT GGC TGC GGT ACT CGG 2139
Gly Thr His Phe Val Leu Glu Ser Pro Leu Asn Gly Cys Gly Thr Arg
495 500 505

CCC CGG TGG TCA GCC CTT GAT GGT GTG GTC TAC TAT AAC TCC ATT GTG 2187
Pro Arg Trp Ser Ala Leu Asp Gly Val Val Tyr Tyr Asn Ser Ile Val
510 515 520

ATA CAG GTT CCA GCC CTT GGG GAC AGT AGT GGT TGG CCA GAT GGT TAT 2235
Ile Gln Val Pro Ala Leu Gly Asp Ser Ser Gly Trp Pro Asp Gly Tyr
525 530 535

GAA GAT CTG GAG TCA GGT GAT AAT GGA TTT CCG GGA GAT ATG GAT GAA 2283
Glu Asp Leu Glu Ser Gly Asp Asn Gly Phe Pro Gly Asp Met Asp Glu
540 545 550

GGA GAT GCT TCC CTG TTC ACC CGA CCT GAA ATC GTG GTG TTT AAT TGC 2331
Gly Asp Ala Ser Leu Phe Thr Arg Pro Glu Ile Val Val Phe Asn Cys
555 560 565 570

AGC CTT CAG CAG GTG AGG AAC CCC AGC AGC TTC CAG GAA CAG CCC CAC 2379
Ser Leu Gln Gln Val Arg Asn Pro Ser Ser Phe Gln Glu Gln Pro His
575 580 585

FIG. 10C

GGA AAC ATC ACC TTC AAC ATG GAG CTA TAC AAC ACT GAC CTC TTT TTG 2427

Gly Asn Ile Thr Phe Asn Met Glu Leu Tyr Asn Thr Asp Leu Phe Leu

590 595 600

GTG CCC TCC CAG GGC GTC TTC TCT GTG CCA GAG AAT GGA CAC GTT TAT 2475

Val Pro Ser Gln Gly Val Phe Ser Val Pro Glu Asn Gly His Val Tyr

605 610 615

GTT GAG GTA TCT GTT ACT AAG GCT GAA CAA GAA CTG GGA TTT GCC ATC 2523

Val Glu Val Ser Val Thr Lys Ala Glu Gln Glu Leu Gly Phe Ala Ile

620 625 630

CAA ACG TGC TTT ATC TCT CCA TAT TCG AAC CCT GAT AGG ATG TCT CAT 2571

Gln Thr Cys Phe Ile Ser Pro Tyr Ser Asn Pro Asp Arg Met Ser His

635 640 645 650

TAC ACC ATT ATT GAG AAT ATT TGT CCT AAA GAT GAA TCT GTG AAA TTC 2619

Tyr Thr Ile Ile Glu Asn Ile Cys Pro Lys Asp Glu Ser Val Lys Phe

655 660 665

TAC AGT CCC AAG AGA GTG CAC TTC CCT ATC CCG CAA GCT GAC ATG GAT 2667

Tyr Ser Pro Lys Arg Val His Phe Pro Ile Pro Gln Ala Asp Met Asp

670 675 680

AAG AAG CGA TTC AGC TTT GTC TTC AAG CCT GTC TTC AAC ACC TCA CTG 2715

Lys Lys Arg Phe Ser Phe Val Phe Lys Pro Val Phe Asn Thr Ser Leu

685 690 695

CTC TTT CTA CAG TGT GAG CTG ACG CTG TGT ACG AAG ATG GAG AAG CAC 2763

Leu Phe Leu Gln Cys Glu Leu Thr Leu Cys Thr Lys Met Glu Lys His

700 705 710

CCC CAG AAG TTG CCT AAG TGT GTG CCT CCT GAC GAA GCC TGC ACC TCG 2811

Pro Gln Lys Leu Pro Lys Cys Val Pro Pro Asp Glu Ala Cys Thr Ser

715 720 725 730

CTG GAC GCC TCG ATA ATC TGG GCC ATG ATG CAG AAT AAG AAG ACG TTC 2859

Leu Asp Ala Ser Ile Ile Trp Ala Met Met Gln Asn Lys Lys Thr Phe

735 740 745

ACC AAG CCC CTT GCT GTG ATC CAC CAT GAA GCA GAA TCT AAA GAA AAA 2907

Thr Lys Pro Leu Ala Val Ile His His Glu Ala Glu Ser Lys Glu Lys

750 755 760

GGT CCA AGC ATG AAG GAA CCA AAT CCA ATT TCT CCA CCA ATT TTC CAT 2955

Gly Pro Ser Met Lys Glu Pro Asn Pro Ile Ser Pro Pro Ile Phe His

765 770 775

GGT CTG GAC ACC CTA ACC GTG ATG GGC ATT GCG TTT GCA GCC TTT GTG 3003

Gly Leu Asp Thr Leu Thr Val Met Gly Ile Ala Phe Ala Ala Phe Val

780 785 790

ATC GGA GCA CTC CTG ACG GGG GCC TTG TGG TAC ATC TAT TCT CAC ACA 3051

Ile Gly Ala Leu Leu Thr Gly Ala Leu Trp Tyr Ile Tyr Ser His Thr

795 800 805 810

FIG. 10D

```
GGG GAG ACA GCA GGA AGG CAG CAA GTC CCC ACC TCC CCG CCA GCC TCG  3099
Gly Glu Thr Ala Gly Arg Gln Gln Val Pro Thr Ser Pro Pro Ala Ser
                815                 820                 825

GAA AAC AGC AGT GCT GCC CAC AGC ATC GGC AGC ACG CAG AGC ACG CCT  3147
Glu Asn Ser Ser Ala Ala His Ser Ile Gly Ser Thr Gln Ser Thr Pro
            830                 835                 840

TGC TCC AGC AGC AGC ACG GCC TAGCCCAACC CAGGCCCAAC CCGGCCCAAC     3198
Cys Ser Ser Ser Ser Thr Ala
        845

CCAGCCCAGC CCAGCTCAGC TCAGCTACTC CAAGGGCAGG ACCAATGGCT GAGCCTCGTG  3258
TCCAGACTCA GAGGGCTGGA TTTTGGTTCC CTTGTAAAGA CAGAGTGAAT TTCAGTATAA  3318
AGATCACCCG TTGTATTCAC CCCACACCCA GGGCTAGTAT AAACATGACC CTGGGCTTCT  3378
GTACCACACT AGAATTCATG TGAGAAAGCT AAAATGGTGG TCTTCTCCAC CAGCCCCTCA  3438
CAGGCTTGGG GGTTTTCTAT GTGAAACACA TGCCAGTTTT TAAAATGCTG CTTTGTCCAG  3498
GTGAGAACAT CCATAATTTG GGGCCCTGAG TTTTACCCAG ACTCAAGGAG TTGGTAAAGG  3558
GTTAATAGCC AGATAGTAGA ACCAGTGAGG AGATGCGGCC AAAGATTCTT TATATCTGAA  3618
CCAAGATGTA AAACAAGAAA TGCTTTGAGG CTTTCTAAGC GATCCTCCTG TCTAATTTGC  3678
ACCTTTGTCT GGATGCACTC TTCTGACCTT GCTGCCACAA CCTGTGGGGT CTGATGTGTC  3738
CCAAGATGGG TGCTGCCCTC AGGGACTGCA CCCTGACAAG TGTTAAGGCA ACATTCCTTG  3798
CTTGTGCCCT GGGCCAAAAC CAATGCTGAT GACCTTATCA GCTTCCTGTT TCTTCCCATA  3858
CTGCATACAC CACTGCAAAA TGTCTTAATG CAAATTTTGT ATTTCTTACA GGCCTACAGA  3918
AATTGAAAAT GACCAAAATC AGGAACCACA GATTTGTGCC CATTCCTAAT ATTTTGTTCT  3978
GCAAATTAAT GTATAATTTG AGGTGAAATT CAGTTATAAA GTCAAGGACG AATTTGCACA  4038
GTGATATATT TCTATGTGTA TGCAAGTACA AGTATATAAT ATGTCACCTG GCACATTCAT  4098
TTTCTCAGTT GAAGAAGAGA AAATTTGAAA ATGTCCTTAT GCTTTTAGAG TTGCAACTTA  4158
AGTATATTTG GTAGGGTGAG TGTTTCCACT CAAAATATGT CAACTTAAAA AAAAA       4213
```

FIG. 10E

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 94/11648

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/71 C07K14/495 C07K16/28 C07K16/42
G01N33/68 A61K38/17 A61K39/395 C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	JOURNAL OF BIOLOGICAL CHEMISTRY., vol.268, no.30, 25 October 1993, BALTIMORE US pages 22710 - 22715 FUKUSHIMA D;BUTZOW R;HILDEBRAND A;RUOSLAHTI E; 'Localization of transforming growth factor beta binding site in betaglycan. Comparison with small extracellular matrix proteoglycans.' see the whole document ---	1-33
X	WO,A,93 10215 (MEMORILA SLOAN-KETTERING CANCER CENTER, US) 27 May 1993 see the whole document ---	1-33
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
12 January 1995	06. 02. 95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Nauche, S

Form PCT/ISA/210 (second sheet) (July 1992)

Intern. ,ial Application No
PCT/US 94/11648

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MOLECUCLAR REPRODUCTION AND DEVELOPMENT, vol.32, no.2, June 1992 pages 99 - 104 MASSAGUÉ, J. ET AL.; 'TGF-beta receptors' see the whole document	1-18, 22-25

1

..ernational application No.

PCT/US94/11648

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 21, 26-27, 29, 30, 32, 33 are directed to a method of treatment of the human/animal body the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. [ ] As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

…formation on patent family members

Intern…al Application No
PCT/US 94/11648

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9310215	27-05-93	AU-A- 3178893	15-06-93